Universidade de Aveiro
Ano 2018

Departamento de Física

Jaime Manuel Maia Andrade

Uma análise crítica do modelo de condução de Contacto de Ponta Quântico em memórias com comutação resistiva baseadas em Ta_2O_5

A critical analysis of the Quantum Point Contact model of conduction in Ta_2O_5-based resistive switching memories

Universidade de Aveiro
Ano 2018

Departamento de Física

Jaime Manuel Maia Andrade

Uma análise crítica do modelo de condução de Contacto de Ponta Quântico em memórias com comutação resistiva baseadas em Ta_2O_5

A critical analysis of the Quantum Point Contact model of conduction in Ta_2O_5-based resistive switching memories

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Física, realizada sob a orientação científica do Prof. Dr. Nikolai Andreevitch Sobolev, Professor Associado do Departamento de Física da Universidade de Aveiro

À minha família e amigos, que me acompanharam ao longo deste percurso e ajudaram a chegar onde estou.

o júri

presidente

Prof. Dra. Margarida Maria Resende Vieira Facão
Professora Auxiliar do Departamento de Física da Universidade de Aveiro

vogais

Prof. Dr. Mikhail Igorevich Vasilevskiy
Professor Catedrático do Departamento de Física da Universidade do Minho (arguente)

Prof. Dr. Nikolai Andreevitch Sobolev
Professor Associado do Departamento de Física da Universidade de Aveiro (orientador)

agradecimentos

Agradeço ao Professor Nikolai Sobolev pela orientação científica que me forneceu, e oportunidades que me proporcionou. Agradeço ainda ao Carlos Rosário, pela interminável ajuda dentro e fora do laboratório, sem a qual esta tese não teria sido possível.

Agradeço ao Departamento de Física da Universidade de Aveiro e ao laboratório associado I3N pelas condições necessárias à realização das experiências aqui descritas. Agradeço ainda ao Departamento pelos cinco anos de formação científica que me proporcionou, e pelas amizades que consegui fazer ao longo desse tempo, e ao grupo do *Institut für Werkstoffe der Elektrotechnik 2* da Universidade de Aachen pelo crescimento dos dispositivos.

Agradeço aos técnicos do Departamento, Ivo Mateus, Miguel Rocha, Júlio Gonçalves e José Januário pela disponibilidade e auxílio na componente experimental. Agradeço ainda ao David Furtado pela ajuda no pedido de instalação da terra, e ao Paulo Gonçalves pelo auxílio no *bonding* e na construção do *sample holder*.

Agradeço ao Nuno Caçoilo pela companhia e apoio, e ao Bruno Teixeira pelo auxílio quando necessário. Agradeceria a mais pessoas em particular, mas para não deixar ninguém de parte, ninguém será incluído, e como tal vou incluir-vos a todos. Um obrigado geral aos Zés, Jorges, Anas, Joanas e *et als.* Vós sabeis quem sois e o que fizestes, e se não sabeis ficastes a saber. Aveiro é vosso.

Nós somos Engenharia Física. O curso fala por nós.

palavras-chave

Comutação resistiva, Ta_2O_5, filmes finos de óxidos, sistemas memristivos, contacto de ponta quântico, mecanismos de condução

resumo

As memórias resistivas de acesso aleatório baseadas em redox (*redox-based resistive random access memories*, ou ReRAMs) são candidatas promissoras para implementar uma nova classe de memórias, denominadas memórias de classe de armazenamento (*storage class memories*, ou SCMs). Estas destinam-se a alcançar baixos tempos de latência, a um preço acessível, encaixando-se entre as memórias Flash e as RAMs dinâmicas (*dynamic RAMs*, ou DRAMs). ReRAMs também estão a ser aplicadas no campo de pesquisa de redes neurais, dada a sua capacidade de emular a plasticidade sinática. Como tal, há um interesse crescente em estudar esta classe de dispositivos, conhecidos como dispositivos memristivos, ou com comutação resistiva (*resistive switching*, ou RS).

Este trabalho foca-se nos mecanismos de condução propostos para modelar a corrente elétrica em células com RS. Um em particular, denominado contacto de ponta quântico (*quantum point contact*, ou QPC), foi estudado em profundidade. Para tal, um dispositivo memristivo de Pt/Ta/Ta_2O_5/Pt foi estudado, e curvas de corrente-tensão ($I - V$) obtidas para ambos os estados de resistência. Isto foi repetido para vários valores da conformidade de corrente aplicada. Um método matemático para aplicação do modelo QPC foi então desenvolvido, envolvendo o uso dos algoritmos de Gauss-Newton e Levenberg-Marquardt. Este último foi utilizado com uma abordagem heurística para o peso de regularização, de forma a facilitar a sua aplicação em massa. As taxas de convergência, influência dos parâmetros iniciais e adequação do ajuste foram todas medidas e contabilizadas no desenvolvimento desta abordagem.

Duas aproximações deste modelo foram consideradas. Na primeira, apenas a primeira subbanda no canal condutor contribui para a condução. Na segunda, a altura da barreira é fixa, para além da primeira aproximação. Determinou-se que o modelo original era difícil de aplicar: os parâmetros iniciais apresentaram uma grande influência nos resultados do ajuste, e o algoritmo não foi robusto. A primeira aproximação foi capaz de fornecer bons ajustes aos dados, e de fazê-lo melhor do que os outros mecanismos de condução considerados. Contudo, a sua base física foi criticada, e certas considerações na interpretação dos resultados devem ser tomadas. Argumentou-se contra a segunda aproximação. Esta foi capaz de fornecer ajustes adequados aos dados experimentais, mas a evolução dos parâmetros não correspondeu às previsões do modelo.

keywords

Resistive switching, Ta_2O_5, thin oxide films, memristive systems, quantum point contact, conduction mechanisms

abstract

Redox-based resistive random access memories (ReRAMs) are promising candidates to implement a new class of memories, called storage class memories (SCMs). These are meant to achieve small latency times, at an affordable price, fitting in between Flash memories and dynamic RAMs (DRAMs). ReRAMs are also being applied in the neural network field of research, given their ability to emulate synaptic plasticity.Thus, there is a growing interest in studying this class of devices, known as memristive, or resistive switching, devices.

This work focuses on the conduction mechanisms proposed to model the electrical current in RS devices. One in particular, called quantum point contact (QPC), was studied in depth. With this intent, a Pt/Ta/Ta_2O_5/Pt memrisive device was studied, and current-voltage ($I-V$) curves for both resistance states obtained. This was repeated for various values of applied current compliance. A mathematical method for applying the QPC model was then developed, involving the Gauss-Newton and Levenberg-Marquardt algorithms. The latter was used with a heuristic approach to the regularization weight, to facilitate its application *en masse*. The convergence rates, influence of starting parameters and goodness of fit were all measured and accounted for in developing this approach. Two approximations of this model were considered. In the first, only the first subband in the conducting channel contributes to the conduction. In the second, the barrier height is fixed, in addition to the first approximation. The original model was found to be hard to apply: the starting parameters had a large influence on the fitting results, and the algorithm was not robust. The first approximation was able to provide good fits to the data, and to do so better than the other conduction mechanisms considered. However, its physical basis was criticized, and certain considerations in interpreting the results must be taken. The second approximation was argued against. It was able to provide adequate fits to the experimental data, but the parameters' evolution failed to match the model's predictions.

Index

Index i
List of Figures ii
List of Tables vi
Nomenclature vii
Introduction 1
1 State of the Art 2
1.1 Emerging memories and current status 2
1.2 Memristive systems 3
1.3 Resistive switching 5
1.4 Resistive switching mechanisms 6
1.5 Typical conduction mechanisms 9
2 Quantum Point Contact Model 13
2.1 Theoretical basis 13
2.2 Single subband approximation 15
2.3 Constant barrier height approximation 16
3 Experimental Details and Methodologies Used 17
3.1 Sample description 17
3.2 Equipment and experimental methodologies used 17
3.3 Algorithmic methodologies used 21
4 Results and Discussion 24
4.1 Resistive switching observed 24
4.2 Convergence and guess influence in the QPC model's application 27
4.3 Comparison between the QPC and conventional models 29
4.4 Temperature measurements and observed RTS 34
4.5 Results of the single subband approximation 37
4.6 Results of the fixed barrier height approximation 42
Conclusions and Future Work 46
References 47

List of Figures

1.1 Diagram of the six relationships that have been established between the four fundamental circuit variables: voltage V, charge q, flux φ and current I. Two of them are given by the conservation of charge (between q and I) and by Faraday's Law of Induction (φ and V). Three constitute the usual basic circuit elements, the resistor (between V and I), the capacitor (q and V) and the inductor (φ and I). The remaining relationship between φ and q is thus provided by the memristor. 4

1.2 (a, b) $I-V$ curve for unipolar (a) and bipolar (b) switching, selected for in the same device by applying a different CC (grey dashed line) during the forming step. The voltage for RESET in the unipolar case can be seen to be generally superior to that in the bipolar case. The virgin state can also be seen to be much more resistive that the HRS later achieved during switching. Taken from [48]; (c) $I-V$ curve for a complementary resistive switching device. The different states of the two stacked cells (a black triangle means LRS, none means HRS) give the overall state of the device. Adapted from [53]. 6

1.3 Schematic $I-V$ curve, and structural representation of the filament's evolution, for a VCM device. The green spheres represent oxygen vacancies, and the purple spheres represent metallic ions in a low valence state. Starting in the OFF state (A), the SET process (B) introduces the vacancies into the disc, leading to the ON state (C) being achieved. An opposite voltage polarity induces the reverse migration of vacancies (D), leading back to the OFF state. To the right is a sketch delineating the plug and disk components of the filament. Adapted from [66]. 9

1.4 Energy band diagram displaying: (1) Schottky emission; (2) P-F emission; (3) NNH; (4) Direct tunnelling. ϕ_C and ϕ_A are the energy barrier heights at the cathode and anode interfaces respectively; E_{FC} e E_{FA} the Fermi level energies for both electrodes; V the applied voltage; E_C the minimum of the conduction band of the oxide layer; E_V the maximum of the valence band of the oxide layer; and E_T the trap energy. These four mechanisms have been chosen to highlight the different ways through which conduction can occur in the device, including two possibilities for conduction via traps: tunnelling (3) and thermal excitation (2). 11

2.1 (Left) Diagram of the saddle potential around the constriction point; (Right) One can approximate this saddle potential as a parabolic potential well (red curve), which defines the quantized energy levels E_n along the y-direction (orthogonal to the filament), and a parabolic potential barrier along the x-direction (blue curve), hampering transmission across the filament. Taken from [90]. 14

2.2 (Up) Diagram of the evolution of the constriction between the LRS (lower barrier height) and the HRS (larger barrier height); (Down) Energy diagram of the resulting barrier for both states, highlighting the difference in height (Φ). Here t_B is the width of the barrier at the Fermi level – another measure for α, V is the total applied voltage, $V-IR$ the voltage drop across the device, and β and $(1-\beta)$ the 16

fractions of voltage that drop at the cathode and anode respectively. Adapted from [78, 81].

3.1 (Left) Schematic representation of the sample's layer stacks and geometry (the 1 pads connect to the BE and the 2 to the TE). Devices with different areas were used, in the range from 5^2 to 25^2 μm^2. Taken from [94]. (Right) Schematic representation of the filament's proposed geometry, based on the VCM model previously introduced, highlighting the location of the constriction as per the QPC model. 17

3.2 (Left) Open cryostat, displaying the vacuum shroud (a), radiation shield (b) and cold finger (c). A silver gasket (d) is placed between the cold finger and the sample holder, to improve the thermal contact between the two. (Right) Close-up of the top of the cold finger, with the location of the heat source (e) highlighted. The various wires and the thermocouple used are also visible. Teflon tape and string are used to maintain the electrical wires and thermocouple in physical contact with the cold finger, to ensure no contact with the vacuum shroud. 18

3.3 (a) Schematic representation of the printed circuit board designed, with the dimensions displayed in millimetres; (b) Resultant printed circuit board after etching; (c) Final sample holder, with an already bonded sample. The bonded wires were connected to the lower pads closer to the sample, and the upper pads were used to weld contacts. Those were initially meant to be used as connectors, but proved too fragile for a good contact to be achieved through this method. They were instead used as extensions to allow the welded wires to be further away from the bonded wires. While a design with the pads placed further away could have been created and implemented, this adaptation proved sufficient for the purpose at hand. 19

3.4 (Left) back of the sample holder after connection to the cold finger. The position of the thermocouple (a) is highlighted; (Right) Front of the sample holder after welding the electrical wires to the contacts (b). 19

3.5 Micropositioner probes used to test the sample prior to bonding. 20

3.6 (Left) Adaptor connecting the wires inside the cryostat to the coaxial cables, while also lifting the ground connection between the coaxial shield and the cryostat's body. (Right) Metallic box housing the contacts at the Sourcemeter, and the connection of the coaxial shields to the low-noise ground jack. 20

4.1 Experimental $I - V$ curves for the four used values of CC, after accounting for the series resistance. The grey curves represent all 1250 cycles, while the black curve highlights a single representative cycle, that is representative of the norm. A mean cycle was not used on account of the difficulty of calculating an adequate mean cycle with an applied CC. The inset figure is the same plot on a logarithmic scale for the current. 24

4.2 Cumulative probability of the normalized conductance G/G_0 and resistance R, measured at an applied voltage of 0.1 V, for the four values of CC used. The full symbols denote the HRS, and the empty ones the LRS. 25

4.3 Evolution of the normalized conductance (measured at an applied voltage of 0.1 V) with the number of cycles of voltage the sample was subjected to, for the four CC values used. 25

4.4 Cumulative probability of the RESET (left) and SET (right) voltages, for the four CC values used. The SET voltage was here defined as the voltage for which there was a jump of conductance of at least an order of magnitude, or, if such a case did not occur (especially for lower values of CC), the value when the current reached the CC. The RESET voltage was here defined as the voltage for which the current reached its maximum absolute value. 26

4.5 Results of the SCLC fit for a more nonlinear (HRS 1) and linear (HRS 2) HRS curves. The curves were partitioned into two segments, fitting them to $I \propto V^p$ and calculating the resulting powers p_1 and p_2. A residual ${r_p}^2 = (p_1 - 1)^2 + (p_2 - 2)^2$ was minimized to find the best partition. The inset figure highlights the second segment for HRS 2. 31

4.6 Distribution of the indicators of goodness of fit obtained, expressed as $1 - R^2$. These were taken for both device states, for the various considered models, over all cycles taken. The QPC model used was the single subband one. 31

4.7 $I - V$ characteristics obtained in the HRS, for more nonlinear (right) and more linear (left) curves. These were found to be representative of the rest of the data. The lines show the results of the QPC, SMSE/P-F, $I \propto V^p$ and Schottky fits. The values of N, α and Φ are those resultant from the QPC fit, and the value of p is the power of the $I \propto V^p$ fit. 32

4.8 $I - V$ characteristics obtained in the LRS, for more nonlinear (right) and more linear (left) curves. These were found to be representative of the rest of the data. The lines show the results of the QPC, SMSE/P-F, $I \propto V^p$ and Schottky fits. The values of N, α and Φ are those resultant from the QPC fit, and the value of p is the power of the $I \propto V^p$ fit. 33

4.9 Cumulative probability of the values of $r\sqrt{\varepsilon_r}$ obtained from the SMSE/P-F fits of the experimental data. These were taken for both device states, in accordance to equation (4.4) using $d = 5$ nm. 34

4.10 (Left) Evolution of the Fermi-Dirac distribution for different temperatures, versus the energy measured in relation to the Fermi level; (Right) Simulation of the relative difference of the current from its value at 1 K, at 0.1 V, defined as $\delta_I = |I/I_{1\,K} - 1|$. The current was simulated as per equation (2.6), with $N = 1$ and the other parameters taking the values displayed, at an applied voltage of 0.1 V. 35

4.11 Evolution of the sample conductance with cycle number, displaying its abrupt nature. The five levels at $G/G_0 = 1.15$, 1.10, 1.08, 1.035 and 1.00 are highlighted with dashed grey lines. 36

4.12 Experimental data resultant from the 140 curves obtained at $T = 120\ K$. The inset figure contains the overall curves, over the entire range of the applied voltage. The five regions previously introduced were fitted via the single subband QPC 36

approximation, and the resulting curves displayed in colour. The mentioned transitions around $0.5\ V$ are clearly displayed, but multiple transitions between the regions were found along the entire range of applied voltage values.

4.13 Simulation of the transmission using equation (2.4) for the first five subbands, and the resulting total transmission (the contribution of $n > 6$ is neglectable in all scenarios). Four combinations of $\alpha = 2\pi/\hbar\omega_x$ and $\hbar\omega_y$ (related to Φ) were used. 38

4.14 Values of N resulting from the application of the single subband QPC model. They are plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed. 40

4.15 Values of α resulting from the application of the single subband QPC model. They are plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed. 40

4.16 Values of Φ resulting from the application of the single subband QPC model. They are plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed. 41

4.17 Values of Φ vs α obtained. The colour of the dots is tied to $G/G_0/N$ 0.1 V; the colour of the lines is tied to the transmission predicted by (2.4). Both share the same colour scale. 42

4.18 Transmission probability as a function of αΦ considering the approximation made in [143], compared to the one resultant from applying equation (2.4) with $E = 0$. 43

4.19 Values of N resulting from the application of the fixed barrier height approximation, plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed. 45

4.20 Values of t_b resulting from the application of the fixed barrier height approximation, plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed. 45

List of Tables

1.1 Best projected quality parameters for multiple categories of ReRAMs. They are compares with the same parameters for Flash and DRAM memories. In categories where best projected values could not be found, demonstrated ones were displayed instead. Data taken from [5,6,10]. 3

1.2 Current density expression and its electric field and temperature dependency of commonly reported conduction mechanisms in RS devices. Taken from [70,71]. 10

4.1 Values of $n1/n2$ obtained for 18 sets of starting $\left(N/V_0\ (\mathrm{V}), \Phi\ (\mathrm{eV}), \alpha\ (\mathrm{eV}^{-1})\right)$ guesses, for both the original QPC model (where V_0 (V) was used) and the first subband only approximation (where N was used). In the original model, a value of ω_y was used so that the height of the first subband barrier, $\hbar\omega_y \cdot 3/2$, equals the displayed value of Φ. 27

4.2 Values of $\delta = |\mathrm{sol}/\mathrm{ref} - 1| \cdot 100\%$ obtained for 18 sets of starting $\left(N/V_0\ (\mathrm{V}), \Phi\ (\mathrm{eV}), \alpha\ (\mathrm{eV}^{-1})\right)$ guesses, for both the original QPC model and the first subband only approximation. The (0.5, 0.5, 3) and (1, 2, 1) guesses were used as a reference for the original and single subband models, respectively. The dashed boxes indicate cases where convergence was not achieved. 28

4.3 Linearizations performed in order to evaluate goodness of fit of the models presented in Table 1.2. I is the electric current, V the voltage, and A and B constants. 30

Nomenclature

2DEG … *Two-dimensional electron gas*

AE … *Active electrode*

BD … *(Dielectric) Breakdown*

BE … *Bottom electrode*

BL … *Bulk Limited*

BRS … *Bipolar resistive switching*

CC … *Compliance current*

CE … *Counter electrode*

CF … *Conductive filament*

CMOS … *Complementary metal-oxide-semiconductor*

DRAM … *Dynamic random access memory*

ECM … *Electrochemical metallization mechanism/memory*

EL … *Electrode limited*

FeRAM … *Ferroelectric random access Memory*

FRH … *Fixed range hopping*

F-N … *Fowler-Nordheim tunnelling*

GCV … *Generalized cross-validation*

HBD … *Hard breakdown*

HRS … *High resistance state*

IL … *Injection Limited*

LOOCV … *Leave-one-out cross-validation*

LRS … *Low resistance state*

MIEC … *Mixed ionic-electronic conductor*

MIM … *Metal-insulator-metal*

NAND … *NOT-AND*

NNH … *Nearest neighbour hopping*

P-F … *Poole-Frenkel emission*

PCM … *Phase-change memory*

QPC … *Quantum point contact*

RAM … *Random access memory*

ReRAM … *Redox-based resistive random access memory*

RS … *Resistive switching*

RTS … *Random telegraph signal*

SBD … *Soft breakdown*

SCLC … *Space-charge limited conduction*

SCM … *Storage class memory*

STT-MRAM … *Spin-transfer-torque magnetoresistive random access memory*

TAT … *Trap assisted tunnelling*

TCM … *Thermochemical mechanism/memory*

TE … *Top electrode*

TL … *Transport limited*

URS … *Unipolar resistive switching*

VCM … *Valence change mechanism/memory*

VRH … *Variable range hopping*

Introduction

Resistive switching (RS) devices show a change in their two-terminal electrical resistance when subject to a strong electric field. This change is non-volatile, and reversible in nature. Devices with this property naturally lend themselves to the development of memory technologies, giving birth to a class of memories named redox-based resistive random access memories, or ReRAMs. These tend to be applied as storage class memories (SMCs), bridging the gap between the affordability of Flash memories, and the low latency of dynamic RAMs (DRAMs). Other perks of this technology are its easy integration in the CMOS (complementary metal-oxide-semiconductor) dominated market, and use of simple and easily accessible materials, such as semiconductor and transition metal oxides. Along with their other possible applications, such as in neuromorphic circuits, RS devices are an actively studied topic.

One of the areas of investigation in the field of RS devices is that of the conduction mechanisms present. Many have been proposed to explain the conduction in both device states. One in particular, called quantum point contact (QPC) model, has been looked at in this thesis. This model considers that conduction is limited by a constriction in a conductive filament (CF), where conduction is ballistic in nature. It displays the possibility of fitting both resistance states, explaining the differing degrees of nonlinearity observed, and can be compatible with the current view on the RS mechanisms.

The samples investigated in this work were Pt/Ta/Ta_2O_5/Pt stacks. They were subjected to voltage ramps, and their current-voltage ($I-V$) characteristics taken at room temperature. Four different values of applied compliance current (CC) were used. These curves were then fitted as per the different conduction mechanisms considered, including the QPC model, and the findings discussed. To apply the QPC, certain algorithmic considerations had to be taken, which were discussed and analysed as well. The evolution of the current with temperature was also studied. It is hoped that the gathered information, and its discussion, will aid in shedding light into the underlying conduction phenomena in RS devices, and contribute to their current investigation.

This thesis is organized into four chapters. The first constitutes a revision of the current state of the art, covering the state of memories, memristive devices, resistive switching overview and mechanisms, and some common conduction mechanisms in RS devices. The second chapter introduces the QPC model, and the approximations considered and analysed. The third contains the sample description, experimental details, and algorithmic methodologies used. The fourth contains the obtained results, and the discussion on the observations conducted, different fitting models considered, and the application and results stemming from the QPC model. The different ways of applying this model, and the different algorithmic considerations taken in each, were also discussed in this last chapter.

Chapter 1
State of the Art

1.1 Emerging memories and current status

Semiconductor technology has for a long time evolved according to Moore's Law [1], which dictates that the number of transistors in an integrated circuit doubles every two years. The factors involved in the continuous miniaturization of components, such as advances in current lithographic techniques, or the implementation of new cost-effective ones, were elucidated in his later work [2]. However, the limits to miniaturization, be they of a physical nature, growing cost of application, or others, could be seen approaching rapidly [3], and in the modern days have already arrived [4].

When it comes to memory technologies, NOT-AND (NAND) and *dynamic random access memory* (DRAM) technologies dominate the market. The latter is a volatile storage memory, requiring an active power to maintain the stored information. It offers a higher price per Gb of storage, but presents very low latency times. A big challenge for DRAM is that a way to scale beyond the $20\ nm$ feature size is yet to be developed [5,6]. The first, also called Flash memory, provides non-volatile storage at a cheap price per Gb, but offers certain limitations, like high writing voltages and latency times, and low durability [5–7]. Additionally, scaling this technology becomes hard below the $20\ nm$ feature size. Beyond the difficulties in the lithographic process, data retention is challenged by a growing impact of random thermal fluctuations with a reduction in oxide thickness, with problems like reducing retention times and increasing leakage currents becoming severe [8,9].

NAND thus offers affordable storage, with high latency times; DRAM offers low latency, at a higher cost. A new class of memories, named *storage class memory* (SCM), thus emerges, bridging this gap between latency and affordability. To fill this new class of memories, a growing interest in alternative information storage technologies has developed over the last few decades [4,5,7,9–11]. Some the promising candidates are the *spin-transfer-torque magnetoresistive RAM* (STT-MRAM), *ferroelectric RAM* (FeRAM) and *phase-change memory* (PCM). The one this work shall focus on is the *redox-based RAM* (ReRAM) technology. Based on the *resistive switching* (RS) phenomenon, memories based on this technology offer great advantages in areas like feature size, write and erase times and operating voltages (Table 1.1). They also offer the advantage of being CMOS (complementary metal-oxide-semiconductor) compatible, generally not requiring complex materials (RS is regularly achieved with semiconductor and transition metal oxides), and being easily integrated into crossbar structures given their small cell feature size. The greatest challenges with applying this technology into the memory market is a high intrinsic variance, given that the switching phenomenon is stochastic by nature. As such, further work into areas like the device composition is required.

Beyond application in memory technologies, RS has also seen study in the field of neuromorphic computing and neural network development [12–14]. The integration of the computing unit into the memory storage unit allows one to go beyond von Neumann's bottleneck limit: the transit time of information between the processing and memory units is reduced, increasing the effective processing speed [15]. RS devices can be used as synaptic elements, emulating the necessary synaptic plasticity (changing their resistance state based on the applied external input). They also retain their resistance state in a non-volatile fashion, and offer great

area scalability. As such, they present themselves as great candidates for implementation into deep learning and artificial neural networks.

Overall, the RS technology shows promise, but much work still has to be done to profitably introduce it into the market. This work attempts to further explore the phenomenon and shed insight on the intrinsic mechanisms, in order to push deeper into the unexplored frontiers of noveau information technology.

Best projected	**Flash**	**DRAM**	**ReRAM**		
			Electrochemical	**Valence change**	**Thermochemical**
Feature size	$> 10\ nm$	$> 10\ nm$	$< 5\ nm$	$< 5\ nm$	$35\ nm$ has been demonstrated
Access time	$\sim 100\ \mu s$ write, $\sim 10\ \mu s$ read	$> 25\ ns$	$< 1\ ns$	$< 1\ ns$	$< 10\ ns$ has been demonstrated
Retention time	$10\ yr$	$64\ ms$	$> 10\ yr$	$> 10\ yr$	$> 10\ yr$
Write cycles	$\sim 10^5$	$> 10^{16}$	$> 10^{11}$	$> 10^{12}$	$> 10^6$ has been demonstrated
Operation voltage (V)	> 15 write, 5 read	~ 2	< 0.5 write, < 0.2 read	< 1 write, 0.1 read	> 1 write, 0.4 read has been demonstrated
Write energy, cell level (J/bit)	$\sim 10^{-16}$	10^{-14}	10^{-15} has been demonstrated	10^{-16}	Not available

Table 1.1 Best projected quality parameters for multiple categories of ReRAMs. They are compared with the same parameters for Flash and DRAM memories. In categories where best projected values could not be found, demonstrated ones were displayed instead. Data taken from [5,6,10].

1.2 Memristive systems

The resistive switching phenomenon is generally described within the framework of *memristor*. This concept was first introduced by Chua in 1971 [16] as an additional passive two-terminal circuit element based on a proposed missing link (Figure 1.1) between the charge q and the magnetic flux φ. The new quantity, denoted memristance M, is thus defined as

$$M\big(q(t)\big) \equiv \frac{d\varphi\big(q(t)\big)}{dq(t)}. \tag{1.1}$$

One can then take into account the time derivatives of the charge (electric current I, via the conservation of charge) and magnetic flux (voltage V, via the Faraday Induction Law), obtaining the result

$$M\big(q(t)\big) = \frac{\dot{\varphi}\big(q(t)\big)}{\dot{q}(t)} = \frac{V}{I}. \tag{1.2}$$

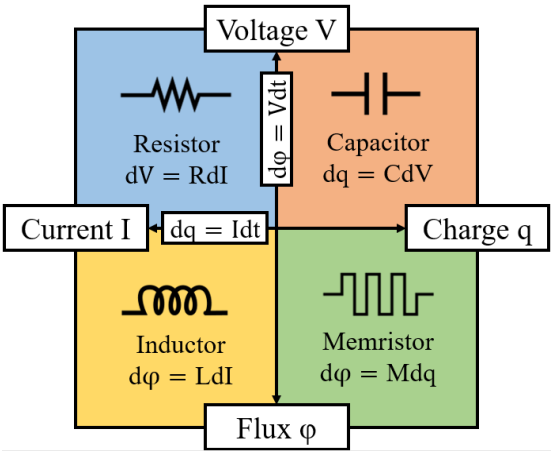

Figure 1.1 Diagram of the six relationships that have been established between the four fundamental circuit variables: voltage V, charge q, flux φ and current I. Two of them are given by the conservation of charge (between q and I) and by Faraday's Law of Induction (φ and V). Three constitute the usual basic circuit elements, the resistor (between V and I), the capacitor (q and V) and the inductor (φ and I). The remaining relationship between φ and q is thus provided by the memristor.

The name *memristor* thus arises from the fact that memristance translates a nonlinear dependence between the voltage and current flowing through the device: it acts as a resistor (note the similarity between (1.2) and Ohm's Law) with a memory of the history of the charge that has flowed through the device. It is thus immediate that any such circuit element could be of interest in the fabrication of devices like memories.

The definition was later broadened into that of *memristive systems* [17], in order to include a class of devices that resemble the memristor, but cannot be realistically modelled by said element alone. Such a class of current controlled devices can be described by

$$V(t) = R(\boldsymbol{x}, I, t)I(t), \tag{1.3}$$

where $\boldsymbol{x}$ is the state of the system (denoting its internal variables that influence the memristance phenomenon, such as magnetization, temperature or a material constant), and

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}, I, t). \tag{1.4}$$

From (1.3) results that $V = 0$ when $I = 0$, meaning an $I - V$ curve of the device will display a pinched hysteresis loop (it crosses the origin). The behaviour of such a curve will depend on the frequency of excitation , such that when the excitation frequency tends to infinity, it tends to a linear resistor [17].

A research team composed by Strukov *et. al.* [18] from HP Labs later proposed that a memristive system could be found in a metal-insulator-metal (MIM) structure based on a TiO_2 thin film. The definition of memristive system has also been used to encompass a myriad of devices, such as other types of memories beyond ReRAMs (like the FeRAMs and PCMs previously mentioned) and thermistors, among others.

Here, the device showed the capacity to change its resistance state when an external voltage was applied, and it was proposed that this was due to ionic transport induced by a strong electric field. However, the qualification of this device within the memristive systems framework, and the framework itself, have been subjects of scrutiny by the scientific community [19–24]. Very relevant criticisms levelled against it include the need for serious corrections in the state equations of the system, such as accounting for the need of Gibbs free energy barriers separating the different resistance states, or the Johnson-Nyquist noise would invariably lead to a loss of information [19,20].

Some care must thus be used with the memristive system label, and in its application to non-volatile resistive switching in particular. Waser [25] proposed that the following considerations should be taken:

- The state variables must include a dependency on the material that constitutes the device, such as its magnetization or crystallographic phase. This avoids the consideration of a purely current controlled device, whose impossibility has commonly been argued for in the cited papers.
- The state variables must have upper and lower boundaries.
- The function $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, I, t)$ must be highly non-linear so as to allow for solutions to the voltage-time dilemma [26].

Even with the constraints and modifications required upon its original definition, the memristive system label is still widely used when it comes to this class of devices, in particular to the resistive switching devices that this thesis focuses on.

1.3 Resistive switching

The resistive switching phenomenon refers to a non-volatile, reversible change in the two-terminal electrical resistance of a device, when subject to a strong electric field. More specifically in the case of ReRAMs, it is considered that nanoionics effects are the underlying cause for the switching: the field-induced motion of said ions results in internal reduction-oxidation (redox) processes, which in turn affect the device's resistance state [25].

Historically, this phenomenon was first described on, and is still widely studied on, MIM structures based on thin transition metal or semiconductor oxide films, such that it is in the oxide film that the ion migration and resistance change is thought to occur [27]. However, RS has also been observed on organic materials and organic-inorganic hybrid compounds [28–33] and on other more exotic materials (like $GaTa_4Se_{8-x}Te_x$ narrow gap Mott insulators [34], $BaTiO_3$ nanocubes [35] and multiferroic heterostructures [36]). Research into new materials that exhibit this phenomenon is currently a hot topic of research.

The device is initially in the virgin state, and a process of electroforming is usually required to begin the switching [37–41]. This process entails applying a high voltage stress to the insulating layer (generally, the forming voltages tend to be higher than the operation voltages) in order to instigate a dielectric breakdown (BD) process. The device is then put into a soft breakdown (SBD) state, where the switching can still be performed. Another possibility is breaking down completely, with the formation of a conductive path across the insulator, in what is known as the hard breakdown (HDB). From this state, switching is no longer possible. In order to prevent this type of breakdown, a limitation on the current flowing through the device, named compliance current (CC), is used both during the forming process, and during the eventual operation of the device. When the current flowing through the device is higher than the specified limit, the voltage source switches to a current source with that value.

Generally, RS devices present two states of resistance, an OFF state with high resistance, the High Resistance State (HRS), and an ON state with low resistance, the Low Resistance State (LRS). It is possible to obtain multiple intermediate resistance states, in what is called multi-state switching [42–46], and which can prove interesting for both high-density memory applications and for neuromorphic computing. However, most applications relate to simple binary computing or memory storage, and the two resistance states are enough to represent the 0 and 1 states of the stored variable in these cases. The transitions between these two states are called SET (High to Low resistance) and RESET (Low to High resistance).

RS can depend on the polarity of the voltage applied: if the SET and RESET operations require opposite polarity applied voltages (Figure 1.2 (b)), it is named *bipolar* resistive switching (BRS).

If both occur for the same polarity (Figure 1.2 (a)), it is named *unipolar* resistive switching (URS). In this last case, the operations differ in the value of the applied voltage that triggers them, with the SET operations usually requiring the higher value. Devices tend to show either one or the other type of switching; however, under certain conditions, such as increasing the forming current or the maximum applied voltage, it has been shown that both can be observed in the same device [47–52].

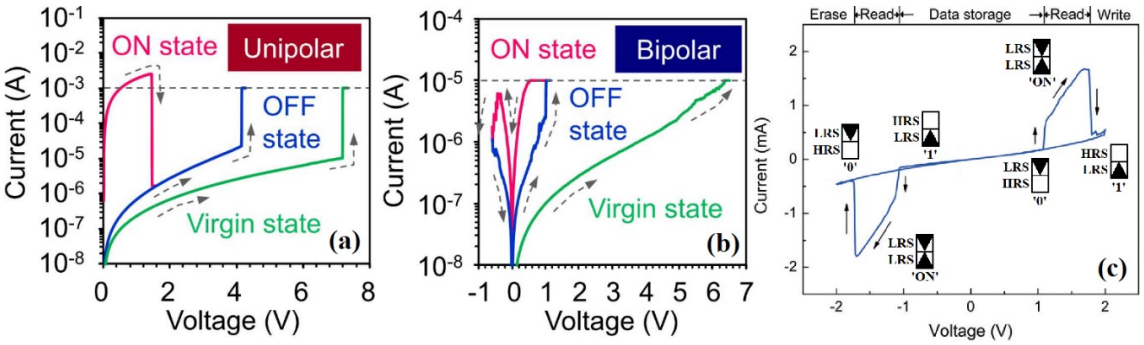

Figure 1.2 (a, b) $I-V$ curve for unipolar (a) and bipolar (b) switching, selected for in the same device by applying a different CC (grey dashed line) during the forming step. The voltage for RESET in the unipolar case can be seen to be generally superior to that in the bipolar case. The virgin state can also be seen to be much more resistive that the HRS later achieved during switching. Taken from [48]; (c) $I-V$ curve for a complementary resistive switching device. The different states of the two stacked cells (a black triangle means LRS, none means HRS) give the overall state of the device. Adapted from [53].

An additional type of RS can be considered, called Complementary Resistive Switching (CRS) [53–58]. This is usually achieved via an anti-serial connection of two resistive switching cells, in what is called a complementary switch: two BRS cells can be stacked anti-serially, so that their respective SET and RESET operations have opposing polarities. One can then reliably set the stack into the ON/OFF and into the OFF/ON states (Figure 1.3), which will be the 0 and 1 states for the device. The advantage of this type of switching is that the overall device will always be in an OFF state, meaning that issues with sneak currents in crossbar arrays are minimized, and scalability is improved (a series transistor or resistor are no longer necessary). However, the reading process is destructive, as it can change the device into an ON/ON state, requiring an additional write operation after the read operation.

This effect can also be achieved in oxide bilayer systems, or even with a single layer [57]. Additionally, due to a competition between electric and thermal processes, it is possible to observe a coexistence of URS and CRS [56], and the evolution of BRS into CRS, attributed to an evolution in a conductive filament asymmetry (more on this in the next section of this chapter), has also been reported [57].

1.4 Resistive switching mechanisms

As mentioned in the previous section, the resistive switching phenomenon has its roots in nanoionics processes taking place in the insulating layer. Currently, it is believed that the conduction in resistive switching devices is filamentary in nature, and the redox reactions taking place lead to the formation and dissolution of one or more conductive filaments (CFs). While interfacial effects can be relevant to the switching phenomenon [59], the commonly accepted

switching models all have in common the creation of a conductive path through the insulator, between the electrodes.

These accepted switching models can be summed up in three categories, based on the nature of the processes that lead to the creation and disruption of the CF: electrochemical, thermochemical and valence-change.

Electrochemical Metallization Mechanism/Memory (ECM)

The ECM [10,60–63], or CBRAM (Conductive Bridge RAM), tends to be applied to MIM structures where one of the electrodes is highly electrochemically active (Active Electrode AE, constituted by metals like Cu, Ag, Ni …), and the other is more inert (Counter Electrode CE, with materials like Pt, Au, W …). The insulator chosen also tends to be a mixed ionic-electronic conductor (MIEC), offering a higher ionic mobility.

The switching typically begins with an electroforming step, where, due to the high applied electric field, the initial anodic dissolution of the AE metal M_{AE} into the insulating layer takes place. This oxidation reaction can be expressed as

$$M_{AE} \rightleftharpoons M_{AE}^{z+} + ze^{-}, \tag{1.5}$$

where M_{AE}^{z+} are the resulting metallic cations of positive charge z and e^{-} the released electrons. Under the influence of the field, these then drift to the CE, where they undergo reduction and electrocristallization reactions, according to the reverse equation,

$$M_{AE}^{z+} + ze^{-} \rightleftharpoons M_{AE}. \tag{1.6}$$

After the electrocristallization process, a metallic filament is formed connecting both electrodes, resulting in a sharp drop in the device's electrical resistance. The application of voltage of a reverse polarity then allows for the dissolution of the filament, as the metallic cations drift back to the AE. As a result, this mechanism of switching always results in BRS.

The electroforming step tends to require higher voltages than the subsequent SET operations. This is due to the initial breaching of the insulating layer, where the first filament's formation creates a structural deformation in the layer that facilitates the subsequent ionic movement through this channel. There also tends to be a high asymmetry in the SET and RESET voltages, as the voltage required to begin the ionic migration and form the filament is smaller than that required to then fully dissolve it.

Thermochemical Mechanism/Memory (TCM)

The TCM [64,65] hinges on thermally activated processes dominating the electrochemical phenomena previously described. In these cells, it is no longer required to use an AE, and both electrodes can contain more electrochemically inert metals, and the insulating layer is usually a transition metal or semiconductor oxide.

The flow of current through the device leads to a local increase in temperature T via Joule heating, according to

$$\sigma E^2 = C_V \frac{\partial T}{\partial t} - \mathrm{div}(\kappa \cdot \mathrm{grad}\, T), \tag{1.7}$$

where σ is the local conductivity, E the applied electric field, C_V the specific heat per volume and κ the local thermal conductivity. Given that the heat is being dissipated into the neighbouring insulating material, the local electric conductivity will then increase, given its dependence on temperature,

$$\sigma \propto \exp\left(-\frac{W_A}{kT}\right), \tag{1.8}$$

where W_A is the activating energy for conduction and k is the Boltzmann constant. Thus, an increase in temperature due to current flow gives rise to an increase in conductivity, increasing the current flow, and the generated heat. In the electroforming process, once the applied field is high enough, one achieves a thermal runaway process, where a thermally induced BD of the insulating layer takes place: given the increase in temperature, the energetically favoured redox reactions can occur, and the added ionic mobility allows for a drifting of oxygen anions out of the high temperature region. This leads to the formation of a metallic filament.

It is believed that the RESET process is also thermally activated, and induced by a reduction in the free surface energy of the CF. This leads to the metallic cations drifting towards the electrodes, opening a gap in the filament. Given that both the SET and RESET processes are dependent on temperature increase, and not on the polarity of the applied field, this switching is strictly unipolar.

The voltage required for the SET operation is lowered after electroforming due to the filament not being totally broken down in the RESET process, facilitating future formation processes. Also, due to the higher voltages required for operating the device, and the fact that RESET is also thermally induced, a CC is applied for the Set operation to limit the current, and heating, of the oxide layer.

Valence Change Mechanism/Memory (VCM)

The VCM [10,66–68] takes place in MIM cells with an AE and CE, and a semiconductor or transition metal oxide as the insulating layer. However, unlike the case for ECM, the AE is not a highly electrochemically active material, but rather one with a low oxygen affinity (Pt, Ir, TiN …). The CE here acts as an Ohmic contact, and a high oxygen affinity, low work function metal is best. Here, the AE provides an interface where the redox reactions occur (it has been shown [69] that it is the ability of this electrode to reduce the semiconducting/metallic species in the oxide layer that is determinant in the choice of AE).

The initial electroforming step breaks down the oxide, generating defects and creating oxide vacancies in the layer. Using the Kröger-Vink notation to represent this process, one obtains

$$O_O \rightleftharpoons \frac{1}{2}O_2 + V_O^{\bullet\bullet} + 2e^-, \tag{1.9}$$

where O_O denotes an oxygen atom in a regular lattice position and $V_O^{\bullet\bullet}$ an oxygen vacancy with double positive charge. The released electrons also change the valence state of the neighbouring metallic atoms, as represented in equation (1.6). This process leads to the formation of a non-switching metallic CF called *plug*, with a switching region located at the interface with the AE called *disc* (Figure 1.3).

The subsequent SET and RESET operations consist in the drift of the generated vacancies through the application of the appropriate voltage polarity. The SET operation attracts the vacancies to the disc, completing the filament and facilitating conduction, while the RESET operation removes them towards the plug, restoring the oxide in the disc, and with it the potential barrier it represents. Given that the writing operations rely on opposite voltage polarities being applied, the switching observed is bipolar, and the SET and RESET voltages tend to be more symmetrical than the ones in the ECM case. The SET voltage also tends to be smaller than the electroforming one, as the plug and disc structure is already formed and only vacancy motion is required, not their creation.

It is possible for these processes to coexist in the same device, as previously mentioned in the coexistence of URS and BRS [47–52]; in these cases, under different conditions one of the underlying mechanisms tends to dominate. For example, an ECM device may exhibit unipolar switching if the applied voltage is driven to high enough values for a thermally induced BD to occur. Here, the device could behave as a TCM memory, until the voltage polarity was inverted, allowing for the ionic migration to become the dominating factor.

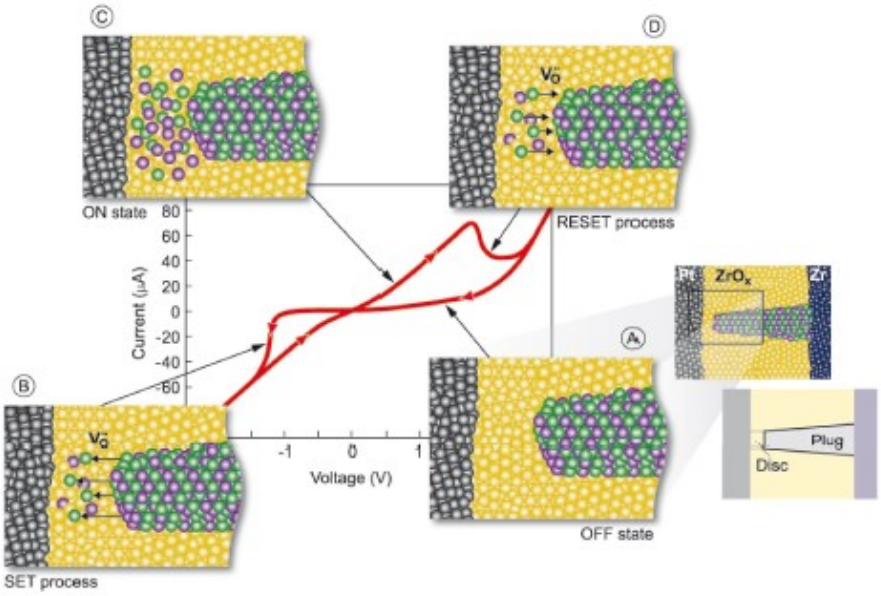

Figure 1.3 Schematic $I - V$ curve, and structural representation of the filament's evolution, for a VCM device. The green spheres represent oxygen vacancies, and the purple spheres represent metallic ions in a low valence state. Starting in the OFF state (A), the SET process (B) introduces the vacancies into the disc, leading to the ON state (C) being achieved. An opposite voltage polarity induces the reverse migration of vacancies (D), leading back to the OFF state. To the right is a sketch delineating the plug and disk components of the filament. Taken from [66].

1.5 Typical conduction mechanisms

Various conduction mechanisms, with differing influences of the electric field and temperature in particular (Table 1.2), have been reported for both the LRS and the HRS in RS devices. Some of those reported including Schottky emission, *Poole-Frenkel* (P-F) emission, *Fowler-Nordheim* (F-N) tunnelling, direct tunneling, *space-charge-limited conduction* (SCLC), Ohmic conduction, hopping (*Mott variable-range hopping*, VRH, and *nearest neighbour hopping*, NNH), *trap-assisted tunnelling* (TAT) and ionic conduction [10,70–72].

These different mechanisms can be classified into two groups: electrode-limited (EL)/injection limited (IL); and transport limited (TL)/bulk limited (BL). The first (EL/IL) denotes mechanisms where it's the properties of the oxide-metal interface, such as the energy barrier heights, that dominate the conduction properties. The mechanisms previously mentioned that fit in this group are: Schottky emission and direct and Fowler-Nordheim tunnelling.

Conduction Mechanism	Current Density Expression[b]	Electric Field and Temperature Dependency[b]
Schottky Emission	$J_{SE} = \frac{4\pi q m^* (kT)^2}{h^3} \exp\left[\frac{-q\left(\phi_B - \sqrt{qE/4\pi\varepsilon}\right)}{kT}\right]$	$J_{SE} \propto T^2 \exp\left(A\frac{\sqrt{E}}{T} - B\right)$
Simmons' modified Schottky equation (SMSE)	$J_{SMSE} = \alpha T^{3/2} E \mu \left(\frac{m^*}{m_0}\right)^{3/2} \exp\left[\frac{-q\left(\phi_B - \sqrt{qE/4\pi\varepsilon}\right)}{kT}\right]$	$J_{SMSE} \propto T^{3/2} E \exp\left(A\frac{\sqrt{E}}{T} - B\right)$
Poole-Frenkel (P-F) Emission	$J_{PF} = q\mu N_c E \exp\left[\frac{-q\left(\phi_B - \sqrt{qE/\pi\varepsilon}\right)}{kT}\right]$	$J_{PF} \propto E \exp\left(A\frac{\sqrt{E}}{T} - B\right)$
Fowler-Nordheim (F-N) Tunnelling	$J_{FN} = \frac{q^2}{8\pi h \phi_B} E^2 \exp\left(\frac{-8\pi\sqrt{2qm^*}}{3hE}\phi_B{}^{3/2}\right)$	$J_{FN} \propto E^2 \exp\left(\frac{-A}{E}\right)$
Direct Tunnelling[c]	$J_{DT} = \frac{q^2}{8\pi h \varepsilon \phi_B} C(V_G, V, t, \phi_B) \cdot \exp\left\{\frac{-8\pi\sqrt{2qm^*}\phi_B{}^{3/2}}{3h\lvert E\rvert}\left[1 - \left(1 - \frac{\lvert V\rvert}{\phi_B}\right)^{3/2}\right]\right\}$	
Space-charge-limited Emission (SCLC)[a]	$J_{SCLC} = \frac{9}{8}\varepsilon_i \mu \theta \frac{V^2}{d^3}$	$J_{SCLC} \propto E^2$
Ohmic Conduction	$J_{Ohmic} = q\mu N_c E \exp\left[\frac{-(E_C - E_F)}{kT}\right]$	$J_{Ohmic} \propto E \exp\left(\frac{-A}{T}\right)$
Nearest Neighbour Hopping (NNH)	$J_{NNH} = \sigma_0 \exp\left(\frac{-T_0}{T}\right) E$	$J_{NNH} \propto E \exp\left(\frac{-A}{T}\right)$
Variable-range Hopping (VRH)	$J_{VRH} = \sigma_0 \exp\left(\frac{-T_0}{T}\right)^{1/4} E$	$J_{VRH} \propto E \exp\left(\frac{-A}{T}\right)^{1/4}$
Trap-assisted Tunnelling (TAT)	$J_{SE} = A \exp\left(\frac{-8\pi\sqrt{2qm^*}}{3hE}\phi_B{}^{3/2}\right)$	$J_{SE} \propto \exp\left(\frac{-A}{E}\right)$
Ionic Conduction	$J_{Ion} = Avr \cdot \exp\left(\frac{-\Delta G^{\neq}}{kT}\right) \cdot \left[\exp\left(\frac{re}{2kT}E\right) - \exp\left(-\frac{re}{2kT}E\right)\right]$	$J_{Ion} \propto \exp\left(\frac{-A}{T}\right) \cdot \left[\exp\left(B\frac{E}{T}\right) - \exp\left(-B\frac{E}{T}\right)\right]$

[a] High field only, low field follows Ohm's Law; [b] A and B are constants; [c] C is a correction function

Table 1.2 Current density expression and its electric field and temperature dependency of commonly reported conduction mechanisms in RS devices. Adapted from [70,71].

Schottky emission: Also called thermionic emission (Figure 1.4), it results from the electrons being thermally excited, and thus able to overcome the potential barrier posed by the oxide layer. Additionally, the image force lowers the barrier at the metal-oxide interface, in what is called the Schottky effect. This conduction mechanism is very commonly observed in dielectric films, especially at very high temperatures, where the thermal excitation of the charge carriers is more impactful.

Simmons' modified Schottky emission: Simmons' equation is more appropriate when the electron's mean free path is smaller than the Schottky barrier width, allowing for scattering events to take place as the electrons pass through the barrier [73,74]. As a consequence, for low applied voltages, the extra linear term dominates conduction, while at high voltages the exponential thermionic component is the main driver of conduction. In this case, bulk effects influence conduction (as scattering becomes significant); thus, SMSE should be classified as both IL and BL.

Direct tunnelling: Tunneling (Figure 1.4) through the barrier arises naturally when approaching the problem from a quantum mechanical point of view. This form of conduction is obviously highly dependent on the thickness of the oxide film and height of the potential barrier.

Fowler-Nordheim tunnelling: Another form of tunnelling occurs when a very high electric field is applied. Under these conditions, the electrons may not see the full thickness of the oxide, but rather a triangular barrier that is thinner, resulting in a higher transmission probability.

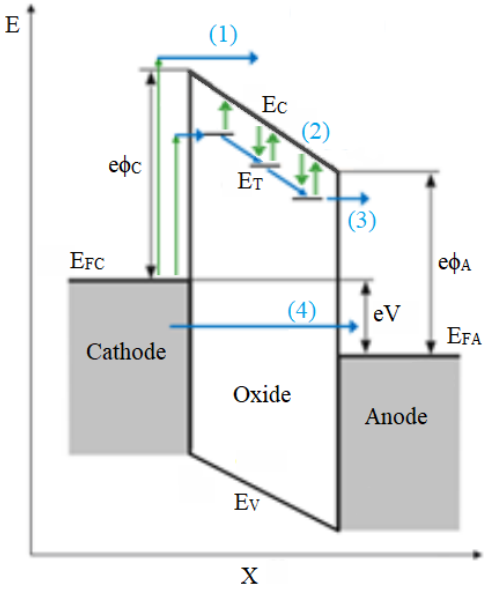

Figure 1.4 Energy band diagram displaying: (1) Schottky emission; (2) P-F emission; (3) NNH; (4) Direct tunnelling. ϕ_C and ϕ_A are the energy barrier heights at the cathode and anode interfaces respectively; E_{FC} e E_{FA} the Fermi level energies for both electrodes; V the applied voltage; E_C the minimum of the conduction band of the oxide layer; E_V the maximum of the valence band of the oxide layer; and E_T the trap energy. These four mechanisms have been chosen to highlight the different ways through which conduction can occur in the device, including two possibilities for conduction via traps: tunnelling (3) and thermal excitation (2).

In the second group (BL/TL), the conduction in the device is controlled by the properties of the oxide layer, such as trap density or electronic mobility. The mechanisms previously mentioned that fit in this group are: Poole-Frenkel emission, SCLC, Ohmic conduction, nearest neighbour and variable range hopping, trap-assisted tunnelling and ionic conduction.

Poole-Frenkel emission: Also called internal Schottky emission (Figure 1.4), it occurs when electrons are thermally excited into the conduction band. However, unlike thermionic emission, the electrons are excited from traps present in the oxide. Here, the trap depth is generally lower than the potential barrier height seen in the Schottky emission. The image charge lowering of the barrier is also doubled, on account of the immobility of the positive charges in this scenario.

SCLC: This mechanism admits three regions of conduction, depending on the intensity of the applied voltage. For low voltages, one is in the linear conduction regime, and $J_{SCLC} \propto E$. Here, the carrier transit time τ_c is higher than the dielectric relaxation time τ_d. Thus, the density of injected carriers is smaller than the density of thermally generated free carriers, and dielectric relaxation occurs (a redistribution of the injected carriers takes place, maintaining the internal charge neutrality in the oxide layer). The linear relation $J_{SCLC} \propto E$ mentioned then results from conduction being dominated by the intrinsic charge carriers (the injected ones are trapped before being able to contribute to conduction).

As the voltage rises, above a threshold voltage V_{th}, one obtains $\tau_c \leq \tau_d$, as the density of injected carriers becomes superior to that of thermally generated ones (the transit time is too short for dielectric relaxation to take place). The injected carriers can now contribute to conduction across the device. Another consequence of the increase in carrier density is the rise of the Fermi level. As this level rises, more traps in the dielectric are filled, until one reaches a second threshold voltage named the trap-filled limit V_{TFL}, where all traps are filled. As the Fermi level rises, the filling of traps in the dielectric creates a space charge filled region in the dielectric, limiting further carrier injection. Child's Law for the current density $J_{SCLC} \propto E^2$ is thus obtained.

Once $V > V_{TFL}$, all traps in the material are filled, and the Fermi level is very close to the bottom of the conduction band. This results in an abrupt increase in the density of available carriers, and also in the current density. Two additional regions of quadratic dependence of J and E can thus be defined, a traps-unfilled SCLC region ($V_{TFL} > V > V_{th}$) and a traps-filled SCLC region ($V > V_{TFL}$).

Ohmic conduction: Here, the mobile charge carriers present in the valence (holes) and conduction (electrons) bands are simply accelerated by the applied electric field, resulting in the familiar linear relation between current density and field $J = \sigma E$. For this mechanism to dominate the conduction across the device, a considerable density of mobile carriers in these bands is required. As such, this mechanism is usually seen in the LRS, and not in the HRS.

NNH: Hopping conduction, much like the P-F emission process mentioned earlier, involves movement of the trapped charge carriers from one trap to the next, facilitating conduction across the device. In P-F emission, this occurs via their thermal excitation into the conduction band; here, the energy of the carriers is not enough to overcome the barrier between the trap level and the bottom of the conduction band, so the movement occurs via tunnelling between the various traps. In the case of NNH (Figure 1.4), also known as *Fixed Range hopping* (FRH), the hopping occurs between the closest spatial neighbours.

VRH: Mott's variable range hopping was suggested to explain conduction in highly disordered systems, where one can find localized charge-carrier states that enable the hopping mechanism. Here, unlike for the NNH case, it is considered that the hopping may occur between traps that are spatially more distant, but where the energy difference between them is lower.

TAT: Here, the tunnelling process is aided by the presence of traps in the dielectric: the electrons are first captured by the trap, and then emitted into the anode. This splits the energy barrier into thinner components, increasing the overall tunnelling probability. This case considers that the traps only aid in tunnelling directly between the electrodes; if tunnelling between multiple traps is considered, one finds oneself with a hopping mechanism (as presented above) instead.

Ionic conduction: This mechanism is based on the drift of ions due to the applied electric field. Given that ions have a much larger mass than electrons, this conduction is not commonly reported in dielectric thin films, but it has been shown to be present in certain resistive switching devices [75].

Chapter 2
Quantum Point Contact Model

2.1 Theoretical basis

Another model that has been considered [76–83] to explain the conduction in RS devices is the *Quantum Point Contact* (QPC) model. This model considers that the CF contains a constriction that, at its narrowest point, is small enough to allow for quantum effects to manifest (it has been reported that the diameter of CFs is of nanometric dimensions [84,85]).This phenomenon is related to that of conductance quantization, first discovered in *2D electron gas* (2DEG) devices [86,87], where it was observed that the conductance of the device increased in a stepwise fashion with the increase of the width of the point contact. While this model was initially used to explore the SBD and HBD behaviour in gate oxide layers [76], it has since then been applied to the switching process as a whole.

The description of this quantization lies in the Landauer-Büttiker formalism [88,89]. In this framework, one considers that:

- The electron mean free path is much greater than the width of the constriction. This results in what is called ballistic conduction, where the electron, not being scattered by impurities, defects or phonons, only changes its motion upon colliding with the walls of the conductor. This is the case for a 2DEG, and, in the case of filamentary conduction, would only be required at the constriction point and not along the entire filament.
- The electrical contacts behave as ideal electron reservoirs, perfectly absorbing all incident electrons and emitting them according to the equilibrium Fermi distribution.

As a result, the constriction can be seen as a potential well, resulting in the quantization of the energy in the direction orthogonal to the filament, as observed in the 2DEG experiments.

One way to model this constriction is to consider the width to change smoothly around the bottleneck (Figure 2.1 left), so that the potential around the point of constriction forms a saddle: the narrowing of the filament around the saddle point claims part of the electronic forward kinetic energy, acting as a potential barrier [81,90,91]. Around the saddle point, both the potential barrier and well can be considered to have a parabolic shape, resulting in the equation for the energy E (and potential V) being

$$E = eV(x,y) = eV_0 - \frac{1}{2}m\omega_x{}^2x^2 + \frac{1}{2}m\omega_y{}^2y^2, \tag{2.1}$$

where V_0 is the value of the potential at the bottleneck (the point $(0,0)$) and ω_x and ω_y frequencies related to the curvature of the saddle along these directions (Figure 2.1 right). As a result of the quantization of energy along the y-direction, which can be described as a function of the quantum number n, (2.1) becomes

$$E_n = eV_0 - \frac{1}{2}m\omega_x{}^2x^2 + \hbar\omega_y\left(n + \frac{1}{2}\right). \tag{2.2}$$

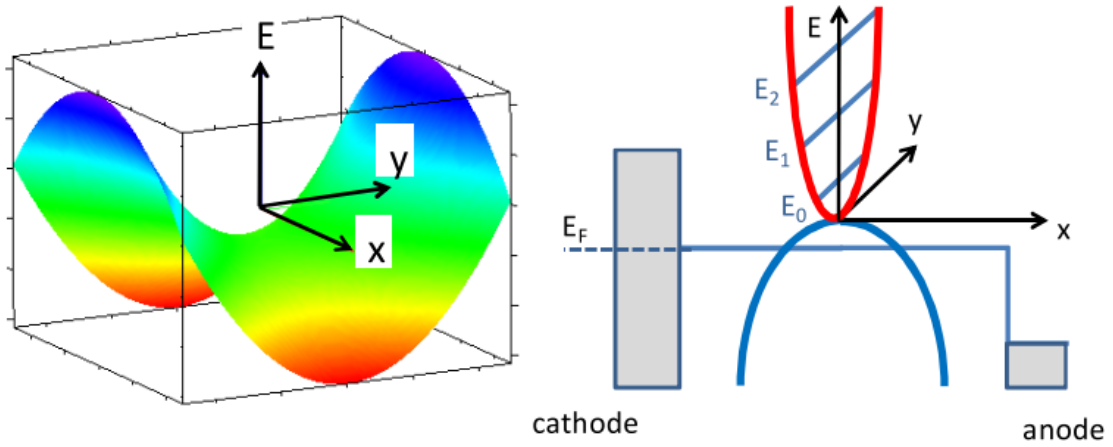

Figure 2.1 (Left) Diagram of the saddle potential around the constriction point; (Right) One can approximate this saddle potential as a parabolic potential well (red curve), which defines the quantized energy levels E_n along the y-direction (orthogonal to the filament), and a parabolic potential barrier along the x-direction (blue curve), hampering transmission across the filament. Taken from [92].

Given the three-dimensionality of the problem, one can assume that the filament displays cylindrical symmetry. Thus, one should not consider a constriction along the y-direction, but rather along the y- and z-directions (or, in cylindrical coordinates, along the r- and θ-directions). The energy term that relates to the quantization thus becomes $\hbar\omega_y(n+1)$. However, as the term appears with the $1/2$ factor in the literature, this is the way it will be used throughout this work.

E_n can be viewed as the band bottom of the nth subband around the bottleneck [90]. The transmission probability between the mth incident subband and the nth outgoing subband, T_{mn}, can then be calculated via the Wentzel–Kramers–Brillouin (WKB) approximation [93,94], resulting in [90]

$$T_{mn} = \delta_{mn}\left\{1+\exp\left[-2\pi\frac{E-\hbar\omega_y\left(n+\frac{1}{2}\right)-eV_0}{\hbar\omega_x}\right]\right\}^{-1}, \tag{2.3}$$

where δ_{mn} is the Dirac delta. As such, transmission only occurs with nonzero probability when the incident and outgoing subbands are the same. One can rewrite (2.3) as

$$T_n = \{1+\exp[\alpha(\Phi_n - E)]\}^{-1}, \tag{2.4}$$

where $\alpha=2\pi/\hbar\omega_x$ is a measure of the curvature of the barrier and $\Phi_n = \hbar\omega_y(n+1/2)+eV_0$ is the height of the barrier for the nth subband. The current I flowing through the filament, from the cathode to the anode, can then be obtained by taking the total transmission, $T=\sum_n T_n$, and solving the integral [81]

$$I = \frac{G_0}{e}\sum_n \int_{-\infty}^{\infty} T_n(E)[f_C(E)-f_A(E)]\mathrm{d}E, \tag{2.5}$$

where $G_0 = 2e^2/h$ is the conductance quantum, and f_C and f_A are the Fermi-Dirac distribution functions for the cathode and anode, respectively. To account for an asymmetrical drop in the potential, (2.5) can be rewritten as

$$I = \frac{G_0}{e}\sum_n \int_{-\infty}^{\infty} T_n(E)\{f_C[E - \beta eV] - f_A[E + (1-\beta)eV]\}\mathrm{d}E, \tag{2.6}$$

assuming that, from the reference point $E = 0$ (the Fermi level for the device at the constriction point, assumed to be constant), the Fermi level of the cathode is βeV above this point, and the Fermi level of the anode is $(1-\beta)eV$ below this point. The voltage across the device V can be expressed as

$$V = V_t - IR, \tag{2.7}$$

where V_t is the total applied voltage and R the series resistance (as such, eV is the total energy difference between both Fermi levels at the respective interfaces). In the expression $(1-\beta)eV$, the term β thus represents the asymmetry in the potential drop, with $\beta = 1/2$ meaning the constriction is halfway along the filament, and $\beta = 0$ or $\beta = 1$ meaning the constriction is close to the interface of one of the electrodes).

The series resistance is thus accounted for (otherwise, (2.6) would be an implicit equation for I), and one assumes that there is no smearing of the Fermi function at the electrodes (zero-temperature limit), such that the applied voltage will drop abruptly at the electrode interfaces with the fractions mentioned above. The integral then takes the form

$$I = \frac{G_0}{e}\sum_n \int_{-(1-\beta)eV}^{\beta eV} T_n(E)\mathrm{d}E. \tag{2.8}$$

This integral then has an analytical solution in

$$I = \frac{G_0}{e}\sum_n \left\{eV + \alpha^{-1}\ln\left[\frac{1+\exp\{\alpha[\Phi_n - \beta eV]\}}{1+\exp\{\alpha[\Phi_n + (1-\beta)eV]\}}\right]\right\}. \tag{2.9}$$

2.2 Single subband approximation

One can stop at equation (2.9) and use it to fit the experimental data, but some research groups [76,78,80] have suggested using an additional approximation, where one considers that only the first subband contributes to the conduction, and then considering that there may be multiple conduction paths (i.e., vacancy paths) within the conductive filament that contribute to conduction. This results in (2.9) taking the form

$$I = \frac{G_0}{e}N\left\{eV + \alpha^{-1}\ln\left[\frac{1+\exp\{\alpha[\Phi - \beta eV]\}}{1+\exp\{\alpha[\Phi + (1-\beta)eV]\}}\right]\right\}, \tag{2.10}$$

where N is the number of such paths and Φ is thus an effective barrier (considering that all parallel paths have the same barrier height and shape). This latter equation immediately benefits from the lack of summation over n, making it faster to implement numerically.

In this interpretation, one stops considering a saddle potential with x and y curvatures, and starts considering that per conduction path only a single barrier exists. As such, for switching between the LRS and HRS, the barrier height and shape change (Figure 2.2): in the former the height is small, and the barrier offering little hindrance to transmission, also resulting in more linear $I - V$ curves, while in the latter it is larger, and the resulting curve is also much more nonlinear.

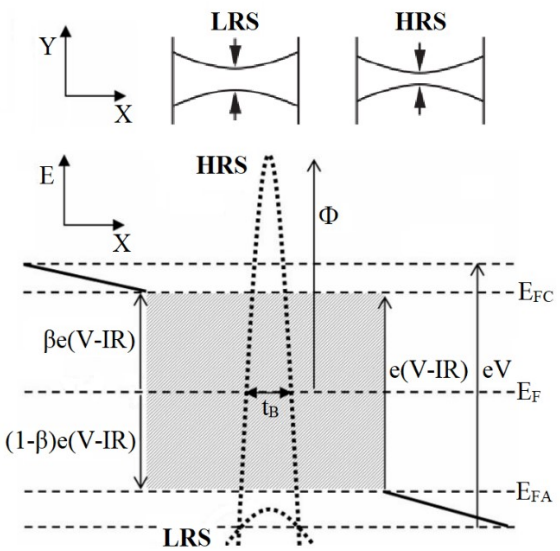

Figure 2.2 (Up) Diagram of the evolution of the constriction between the LRS (lower barrier height) and the HRS (larger barrier height); (Down) Energy diagram of the resulting barrier for both states, highlighting the difference in height (Φ). Here t_B is the width of the barrier at the Fermi level – another measure for α, V is the total applied voltage, $V - IR$ the voltage drop across the device, and β and $(1-\beta)$ the fractions of voltage that drop at the cathode and anode respectively. Adapted from [78,81].

2.3 Fixed barrier height approximation

Another approximation to the model is that one can consider the N conduction paths to be strings of oxygen vacancies, and the barrier is introduced by the removal of a number of vacancies in said path [82]. As per the *ab-initio* calculation in the cited paper, one considers the resulting barrier to have a fixed height, and its width t_b to change (according to the number of oxygen vacancies removed), resulting in the equation

$$I = \frac{G_0}{e} N \left\{ eV + \frac{t_0 \Phi}{t_b} \ln \left[\frac{1 + \exp\left\{ \frac{t_b}{t_0} \left[1 - \beta \frac{eV}{\Phi} \right] \right\}}{1 + \exp\left\{ \frac{t_b}{t_0} \left[1 + (1-\beta) \frac{eV}{\Phi} \right] \right\}} \right] \right\}, \tag{2.11}$$

where t_0 is a parameter resulting from the *ab-initio* calculations. This interpretation has the advantage of possibly removing one parameter from the fit if one uses a value of Φ obtained from said *ab-initio* calculations.

Chapter 3
Experimental Details and Methodologies Used

3.1 Sample description

The devices used (Figure 3.1) were constituted by a Pt/Ta/Ta_2O_5/Pt stack, where the 5 nm thick Ta_2O_5 layer is the active oxide where switching occurs, and the 15 nm thick Ta and 20 nm thick bottom Pt layers are the top electrode (TE) and bottom electrode (BE), respectively (the top 20 nm thick top Pt layer was used to prevent Ta oxidation). The Ta electrode is thought to work as an oxygen scavenger, absorbing oxygen present in the oxide layer and creating additional vacancies [96–99]. Below the bottom Pt layer there is a 3 nm thick Ti layer (to improve adhesion of the Pt layer), a 430 nm thick SiO_2 layer (providing electrical insulation from the substrate), and then the *p*-doped Si substrate.

The samples were grown at the *Institut für Werkstoffe der Elektrotechnik 2* of the RWTH Aachen University, and the growth process was the same as reported in [95].

All layers were deposited via radiofrequency (RF)-magnetron sputtering at room temperature, guaranteeing that the base pressure of the chamber was always lower than 10^{-5} mbar before deposition begins. The Pt films were sputtered in a pure Ar atmosphere, with an operating pressure of 5×10^{-3} mbar and a RF-power of 80 W. The Ta film was sputtered in a pure Ar atmosphere, with an operating pressure of 4×10^{-3} mbar and a RF-power of 100 W. The Ta_2O_5 film was sputtered in a reactive Ar and O_2 atmosphere, in the proportion 3 to 2, with an operating pressure of 4×10^{-3} mbar and a RF-power of 100 W. The samples were then patterned via conventional UV photolithography and liftoff into the crossbar shape depicted in Figure 3.1, with cell areas of 5^2, 10^2 and 25^2 μm^2. The different device areas resulted in different series resistances, with values of approximately 500, 300 and 200 Ω, respectively.

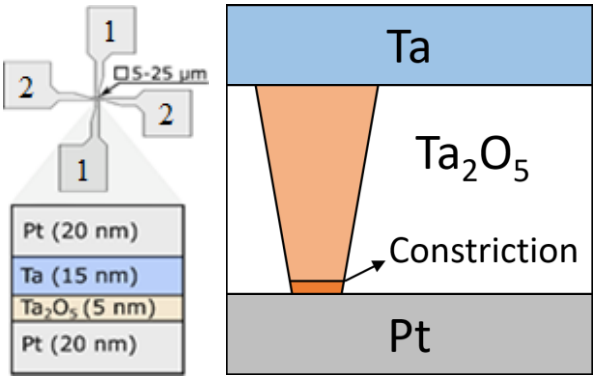

Figure 3.1 (Left) Schematic representation of the sample's layer stacks and geometry (the 1 pads connect to the BE and the 2 to the TE). Devices with different areas were used, in the range from 5^2 to 25^2 μm^2. Taken from [95]. (Right) Schematic representation of the filament's proposed geometry, based on the VCM model previously introduced, highlighting the location of the constriction as per the QPC model.

3.2 Experimental Equipment and Methodologies

To conduct the electrical measurements, a Keithley 2611A Sourcemeter was used to both supply the electrical signals and measure the output. The device was remotely controlled via a program developed in Matlab using some of its built-in commands [100].

The sample was placed inside an ARS CS202-DMX-15 closed cycle cryostat, with a DMX-12 model vacuum shroud and radiation shield (Figure 3.2).

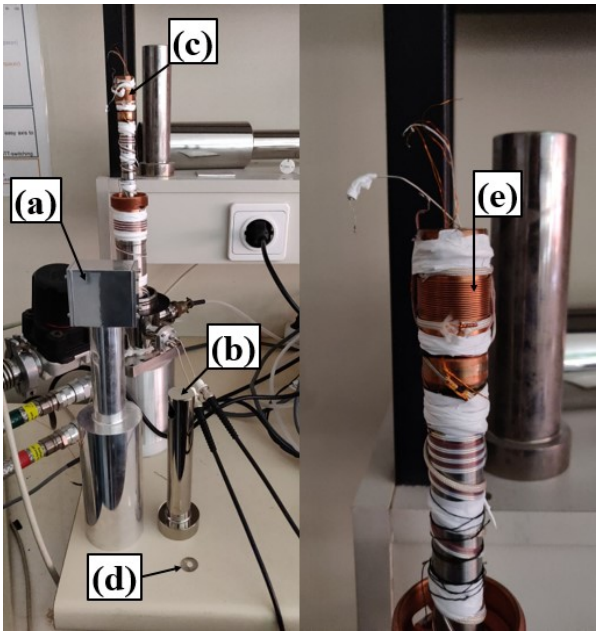

Figure 3.2 (Left) Open cryostat, displaying the vacuum shroud (a), radiation shield (b) and cold finger (c). A silver gasket (d) is placed between the cold finger and the sample holder, to improve the thermal contact between the two. (Right) Close-up of the top of the cold finger, with the location of the heat source (e) highlighted. The various wires and the thermocouple used are also visible. Teflon tape and thin nylon string are used to maintain the electrical wires and thermocouple in physical contact with the cold finger, to ensure no contact with the vacuum shroud.

The sample was placed on the sample holder with the help of a printed circuit board (Figure 3.3), designed to allow the pads to be wire bonded to it, in order to achieve a stable electrical contact that could reach low temperatures (if a simpler configuration with a lower mechanical stability was used, the thermal expansion/contraction of the contacts could have impacts on the quality, or even existence, of the contact). The printed circuit board contained 8 pads for contact with the sample's electrodes, and a metallic extension was welded to the pads, facilitating the welding of the electrical wires present in the cryostat without the risk of accidentally breaking the bonded wires. This printed circuit board was made of a PBC substrate covered by a gold sheet, which was etched into the desired shape. It was then connected to the cryostat's sample holder using screws. The substrate is electrically insulating, and care was taken to ensure that the conductive pads and metallic extensions were distant from the screws, so that the possibility of discharge through the cryostat's cold finger was minimized. The sample itself was glued to the sample holder with carbon tape, in order to ensure that there was a good thermal contact between the two.

A thermocouple (Figure 3.4) was attached to the back of the sample holder, and another to the cold finger, right next to the hot source. The temperature was then measured with a Lake Shore 325 Temperature Controller at both ends, and the temperature of the hot source controlled via the built-in PID algorithm. The thermocouple and wires were enveloped around the cryostat's axis, and fastened using thin nylon string and Teflon tape. A silver gasket was also placed between the cryostat's cold finger and the sample holder, before fastening it, in order to improve the thermal contact.

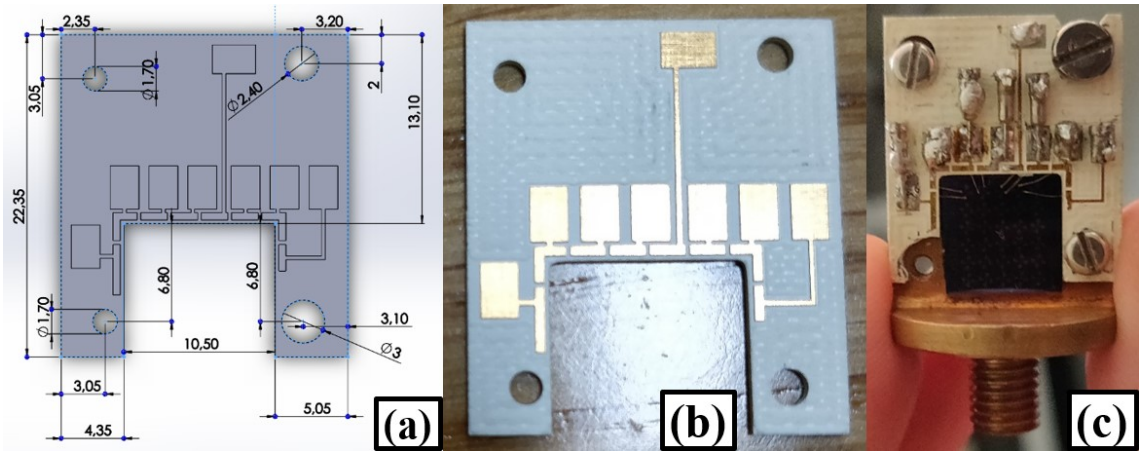

Figure 3.3 (a) Schematic representation of the printed circuit board designed, with the dimensions displayed in millimetres; (b) Resultant printed circuit board after etching; (c) Final sample holder, with an already bonded sample. The bonded wires were connected to the lower pads closer to the sample, and the upper pads were used to weld contacts. Those were initially meant to be used as connectors, but proved too fragile for a good contact to be achieved through this method. They were instead used as extensions to allow the welded wires to be further away from the bonded wires. While a design with the pads placed further away could have been created and implemented, this adaptation proved sufficient for the purpose at hand.

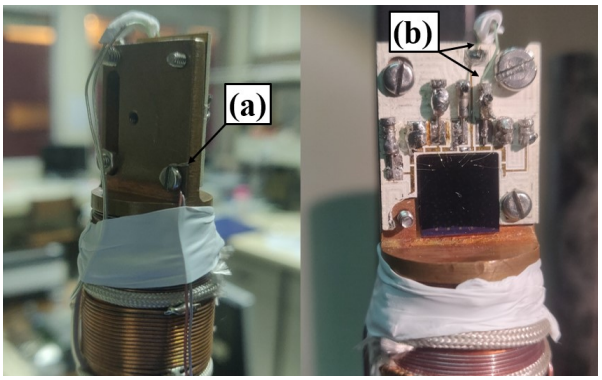

Figure 3.4 (Left) back of the sample holder after connection to the cold finger. The position of the thermocouple (a) is highlighted; (Right) Front of the sample holder after welding the electrical wires to the contacts (b).

Prior to bonding, some of the devices were tested using micropositioner probes (Figure 3.5) in order to identify those which displayed proper switching, and those which had already undergone HBD.

An adaptor (Figure 3.6, left) was created to connect the single wires used inside the cryostat with coaxial cables used outside it. This adaptor disconnected the cryostat chassis (which is connected to the common ground) from the coaxial cable shielding, in order to help prevent ground loops.

A metallic box (Figure 3.6, right) was also used to house the adaptors from coaxial to single wire at the Sourcemeter point of connection. The coaxial shields were then connected to the low-noise chassis ground jack, which is connected to the chassis ground via a frequency-dependent resistor, further insulating the sense channels from high-frequency noise present in the ground.

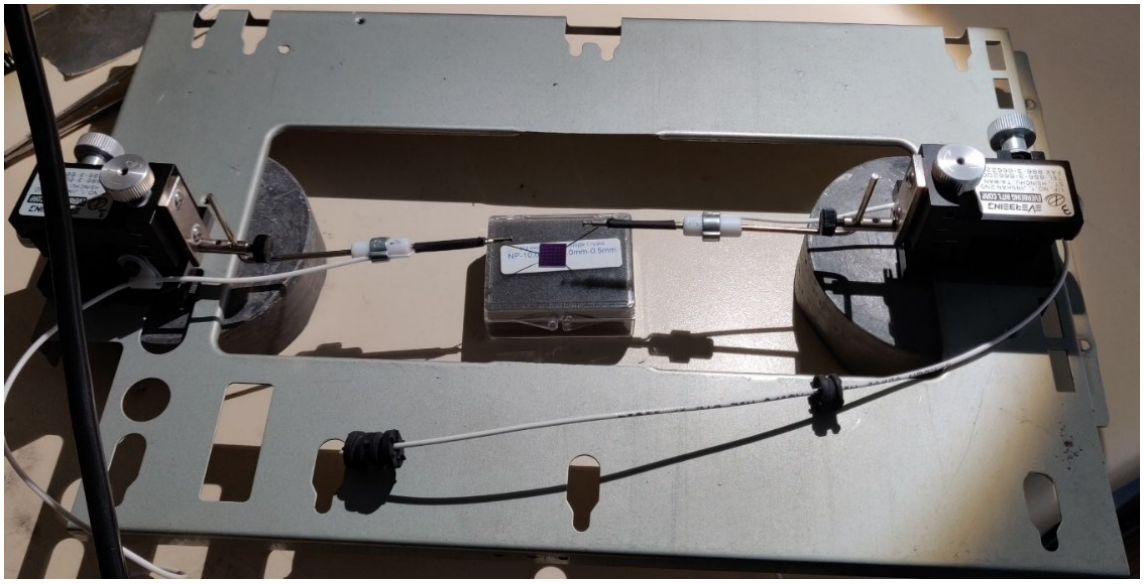

Figure 3.5 Micropositioner probes used to test the sample prior to bonding.

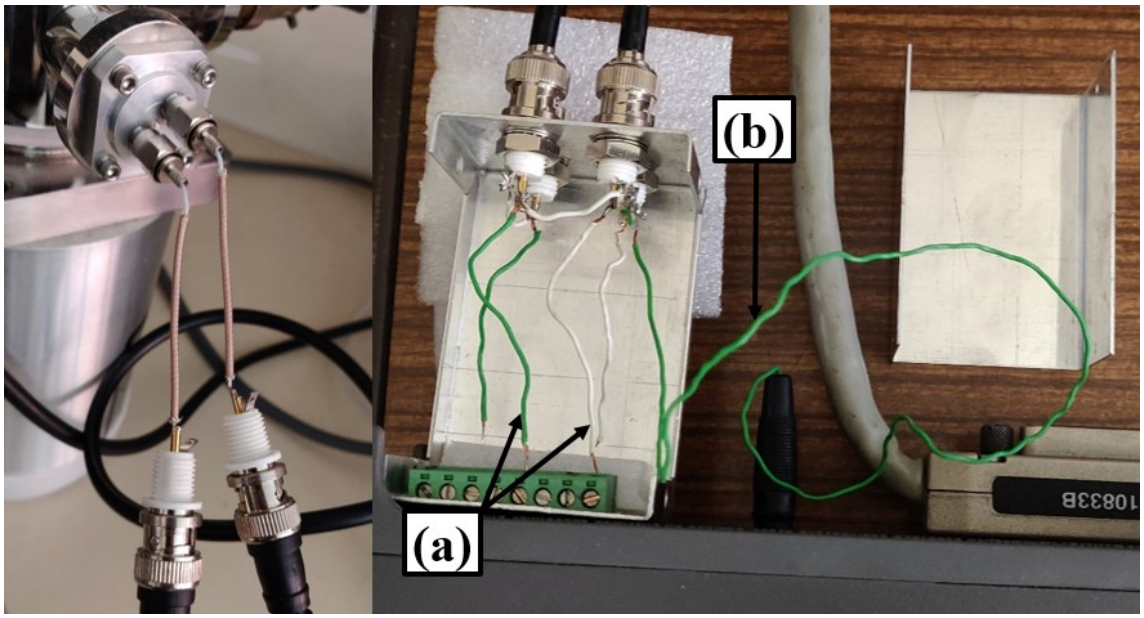

Figure 3.6 (Left) Adaptor connecting the wires inside the cryostat to the coaxial cables, while also lifting the ground connection between the coaxial shield and the cryostat's body. (Right) Metallic box housing the contacts at the Sourcemeter, and the connection of the coaxial shields to the low-noise ground jack.

The devices were subjected to triangular voltage ramps with a rate of approximately $0.3\ \mathrm{V/s}$. The input voltage and output current were measured in a two-wire configuration (only one pair of electrical contacts was used to measure both simultaneously). While this measurement technique is affected by the lead resistances [101], the resistance values found were fairly superior to the $100\ \Omega$ recommended benchmark, and the printed circuit board would have been able to hold only half as many samples simultaneously connected (2 instead of 4). As such, it was determined that a two-wire measurement would suffice.

The CC used in the electroforming step was $50\ \mu\mathrm{A}$.

In order to avoid an over-degradation of the samples, the voltage for the SET operation was limited, so that, after 0.25 V with the current staying at the CC value, the voltage begins to ramp down.

The voltage range was set before measurements began, based on the intended RESET stopping voltage. As the stopping voltage used was -1.5 V, the range selected was 2 V (the smallest value for the range still greater than 1.5 V). On account of the previously mentioned care taken with the SET voltage, this limit was more flexible. The CC and current ranges were also set beforehand, but on occasion changed mid-measurement (for example, when an overshoot was detected at the SET operation). In such cases, these values were changed only when $V = 0$, so as to minimize any possible resultant spikes in the current. Both ranges were as such manually set (autorange was not used).

3.3 Algorithmic Methodologies

The model fitting was performed through the *Gauss-Newton method*, which is an adaption of *Newton's method* in optimization [102–104]. One starts with a multi-variate function f of m variables $\boldsymbol{x} = (x_1, \dots, x_m)$ and n parameters $\boldsymbol{a} = (a_1, \dots, a_n)$, that attempts to fit the experimental data $\boldsymbol{y} = (y_1, \dots, y_m)$. One can then define M residual functions $\boldsymbol{r} = (r_1, \dots, r_M)$to denote the proximity between the fitting and objective functions. For the purposes of this thesis, one shall consider the case of $m = M$, where the residual functions are defined as

$$r_i = y_i - f(x_i, \boldsymbol{a}). \tag{3.1}$$

The sum of squares to be minimized R is thus defined as

$$R = \sum_{i=1}^{m} {r_i}^2 . \tag{3.2}$$

The function R is then, as per the Newton's method [102–104], approximated by its second order Taylor expansion at $\boldsymbol{a} + \Delta\boldsymbol{a}$. Its value, R_T, is

$$R_T = R + \boldsymbol{g} \cdot \Delta\boldsymbol{a} + \frac{1}{2}\boldsymbol{H} \cdot \Delta\boldsymbol{a}^2 + O(\Delta\boldsymbol{a}^3), \tag{3.3}$$

where $\boldsymbol{g}$ is the gradient vector of R and $\boldsymbol{H}$ its Hessian matrix. These can be expressed as functions of the Jacobian matrix $\boldsymbol{J}$ of R, whose elements are

$$J_{ij} = \frac{\partial r_i}{\partial a_j}, \tag{3.4}$$

The elements of $\boldsymbol{g}$ and $\boldsymbol{H}$ are thus given by

$$g_j = 2\sum_{i=1}^{m} r_i \cdot J_{ij}\,, \tag{3.5}$$

$$H_{jk} = 2\sum_{i=1}^{m}\left(J_{ij} \cdot J_{ik} + r_i \cdot \frac{\partial^2 r_i}{\partial a_j \partial a_k}\right). \tag{3.6}$$

One then intends to find an $\boldsymbol{a} + \Delta\boldsymbol{a}$ that is a stationary point of R, so that the derivate of R_T around this point is null. Calculating this from (3.3) gives

$$0 = \frac{\partial}{\partial \Delta \boldsymbol{a}}\left(R + \boldsymbol{g} \cdot \Delta \boldsymbol{a} + \frac{1}{2}\boldsymbol{H} \cdot \Delta \boldsymbol{a}^2\right) = \boldsymbol{g} + \boldsymbol{H} \cdot \Delta \boldsymbol{a}, \tag{3.7}$$

and the resulting recurring relation for Newton's method is

$$\boldsymbol{a}^{k+1} = \boldsymbol{a}^k - \boldsymbol{H}^{-1}\boldsymbol{g}. \tag{3.8}$$

The Gauss-Newton method results from performing a first order approximation rather than a second order one. This can be obtained from Newton's method formalism by ignoring the second order term in the Hessian matrix, so its elements become

$$H_{jk} \approx 2\sum_{i=1}^{m} J_{ij} \cdot J_{ik}\,. \tag{3.9}$$

The main advantage of this method is that calculating the second order terms is thus no longer required, making it faster to apply. Equation (3.8) thus can be simplified into

$$\boldsymbol{a}^{k+1} = \boldsymbol{a}^k - \left(\boldsymbol{J}^{\boldsymbol{T}} \cdot \boldsymbol{J}\right)^{-1} \cdot \boldsymbol{J}^{\boldsymbol{T}} \cdot \boldsymbol{r}. \tag{3.10}$$

One should note that this approximation only holds true if the second order term can truly be neglected. This boils down to ensuring that the bulk of the residuals can be explained by the first order terms alone. This, the so-called small residuals problem [105], thus requires that

$$\left\|\sum_{i=1}^{m} J_{ij} \cdot J_{ik}\right\| \gg \left\|\sum_{i=1}^{m} r_i \cdot \frac{\partial^2 r_i}{\partial a_j \partial a_k}\right\|. \tag{3.11}$$

Issues with the Newton and Gauss-Newton methods arise from two fronts [103,104]. First, the methods are not guaranteed to converge, and in fact they may diverge if the initial guess provided is far from the stationary point. Additionally, the very need of such an initial guess, if the algorithm proves to be overly dependent on it (small deviations leading to divergence), results in difficulties applying the algorithm to big data. Second, if the matrix $\boldsymbol{J}^{\boldsymbol{T}} \cdot \boldsymbol{J}$ is ill-conditioned, one will have issues with amplifying noisy data when attempting to invert it.

A way to minimize these issues comes from applying a regularization method called *the steepest descent* [103,104,106]. Here, this approach was introduced via the *damped least squares method*, or *Levenberg-Marquardt method* [107,108]. This involves adding a term to the Gauss-Newton recurring relation, obtaining

$$\boldsymbol{a}^{k+1} = \boldsymbol{a}^k - \left(\boldsymbol{J}^{\boldsymbol{T}} \cdot \boldsymbol{J} + \lambda \boldsymbol{I}\right)^{-1} \cdot \boldsymbol{J}^{\boldsymbol{T}} \cdot \boldsymbol{r}. \tag{3.12}$$

The additional term $\lambda\boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix and λ is a non-negative scalar, results in adding an additional step to the calculation of the Hessian matrix. This parameter is usually chosen via an *ad-hoc* method, depending on the needs of the specific problem at hand. When $\lambda \to 0$, one obtains the usual Gauss-Newton method; when λ is large, a large step towards the steepest descent of the gradient is taken, and one approaches a pure steepest descent algorithm.

The Levenberg-Marquardt algorithm is thus called a trust-region approach to the Gauss-Newton method, using the added step to improve convergence when far from the solution, with the downside of slowing down convergence when near it. For ill-posed problems, where $\boldsymbol{J}^{\boldsymbol{T}} \cdot \boldsymbol{J}$ is ill-conditioned, it also aids in inverting the overall matrix by damping the influence of noise.

The weight value can be updated throughout the iteration, starting higher in order to more rapidly approach the vicinity of the stationary point desired, and then being lowered to hasten convergence.

Care must be taken to ensure the value of λ is also not too large, as it could prove to be a distortion of the posed problem, and would result in finding local minima far away from the actual global minima one intends to locate. This would exacerbate the issue of dependence upon the initial choice of parameters, rather than lessen it.

The objective becomes to ensure that a good fit can be achieved, while facilitating its application to big data and avoiding an overdependence on the initial guess. The choice of weighting parameter λ was thus evaluated through *leave-one-out cross-validation* (LOOCV) [109,110]. This method of validation splits the n elements sample (here, the experimental I-V point of the curve) into a group with $n-1$ elements, and one with only 1. The model is then trained on the $n-1$ group, and the resulting parameters used to predict the last data point. After repeating this process for each point in the sample, the leave-one-out error rate is approximated via the generalized cross-validation (GCV) statistic,

$$GCV = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{1 - p/n}\right)^2, \tag{3.13}$$

where y_i is the i^{th}outcome of the training set, $\hat{y}_i$ the i^{th} model prediction of said outcome, and p the number of parameters of the model. Given that all testing will be done between models with the same p and n (only the values of λ change between each LOOCV test), one can use a simplified version of the GCV that lies in the usual sum of squared residuals,

$$GCV \to R^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \tag{3.14}$$

The influence of the initial parameters on the solution was also determined, both in relation to the quality of fit, and to the distance between the solutions reached. The latter is important as two different initial guesses could lead to convergence around different local minima, which one would like to avoid. Finding a range of guesses that had little impact on the solution, and provided a good quality of fit, are important when it comes to the application of the model over a large number of individual curves.

Chapter 4
Results and Discussion

4.1 Resistive switching observed

1250 $I-V$ curves were obtained for four values of the current compliance: 50, 100, 200 and 1000 μA. The tested electrodes all had a series resistance value of around 300 Ω, which was accounted for via equation (2.7), before further analysing the curves themselves (Figure 4.1).

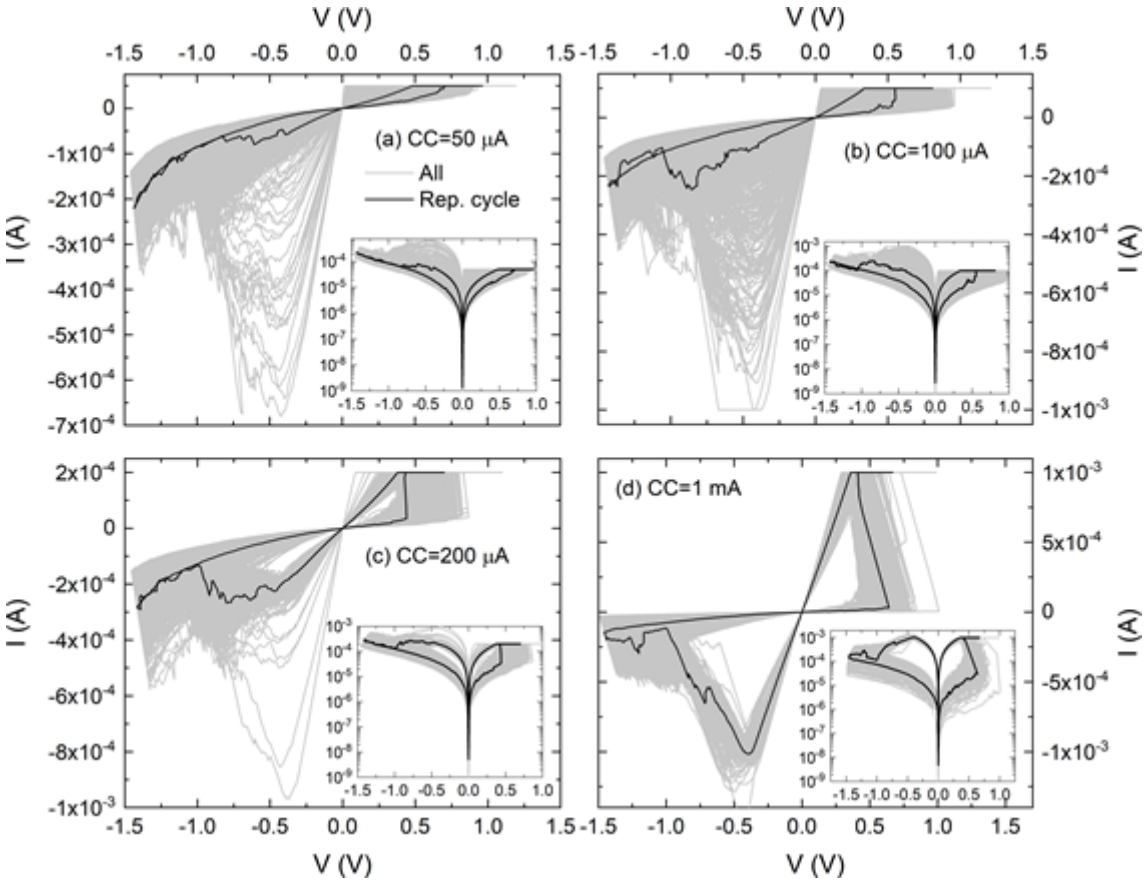

Figure 4.1 Experimental $I-V$ curves for the four used values of CC, after accounting for the series resistance. The grey curves represent all 1250 cycles, while the black curve highlights a single representative cycle, that is representative of the norm. A mean cycle was not used on account of the difficulty of calculating an adequate mean cycle with an applied CC. The inset figure is the same plot on a logarithmic scale for the current.

The observed bipolarity and high symmetry of the SET and RESET operations point to VCM being the underlying switching mechanism present. This is corroborated by the cell's structure, as the Ta TE has a high oxygen affinity (as mentioned, it acts as a scavenging layer) and acts as the Ohmic contact, while the Pt BE, with a higher work function and lower oxygen affinity, acts as the AE where the switching takes place.

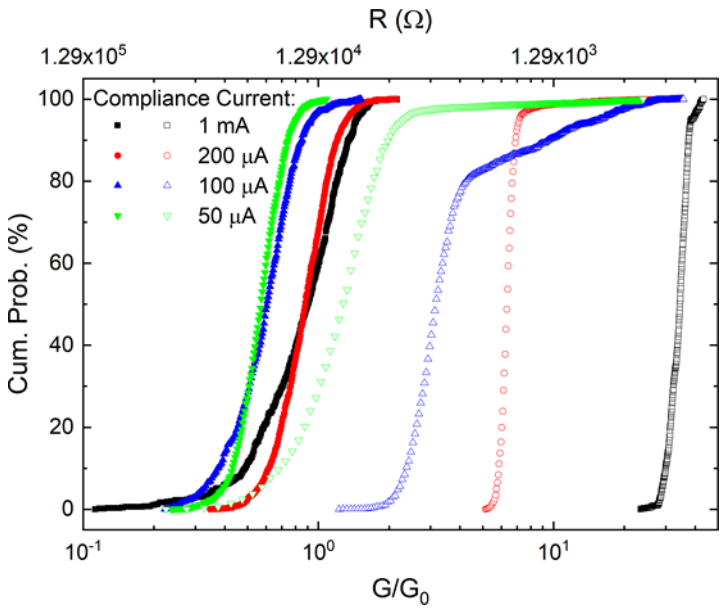

Figure 4.2 Cumulative probability of the normalized conductance G/G_0 and resistance R, measured at an applied voltage of 0.1 V, for the four values of CC used. The full symbols denote the HRS, and the empty ones the LRS.

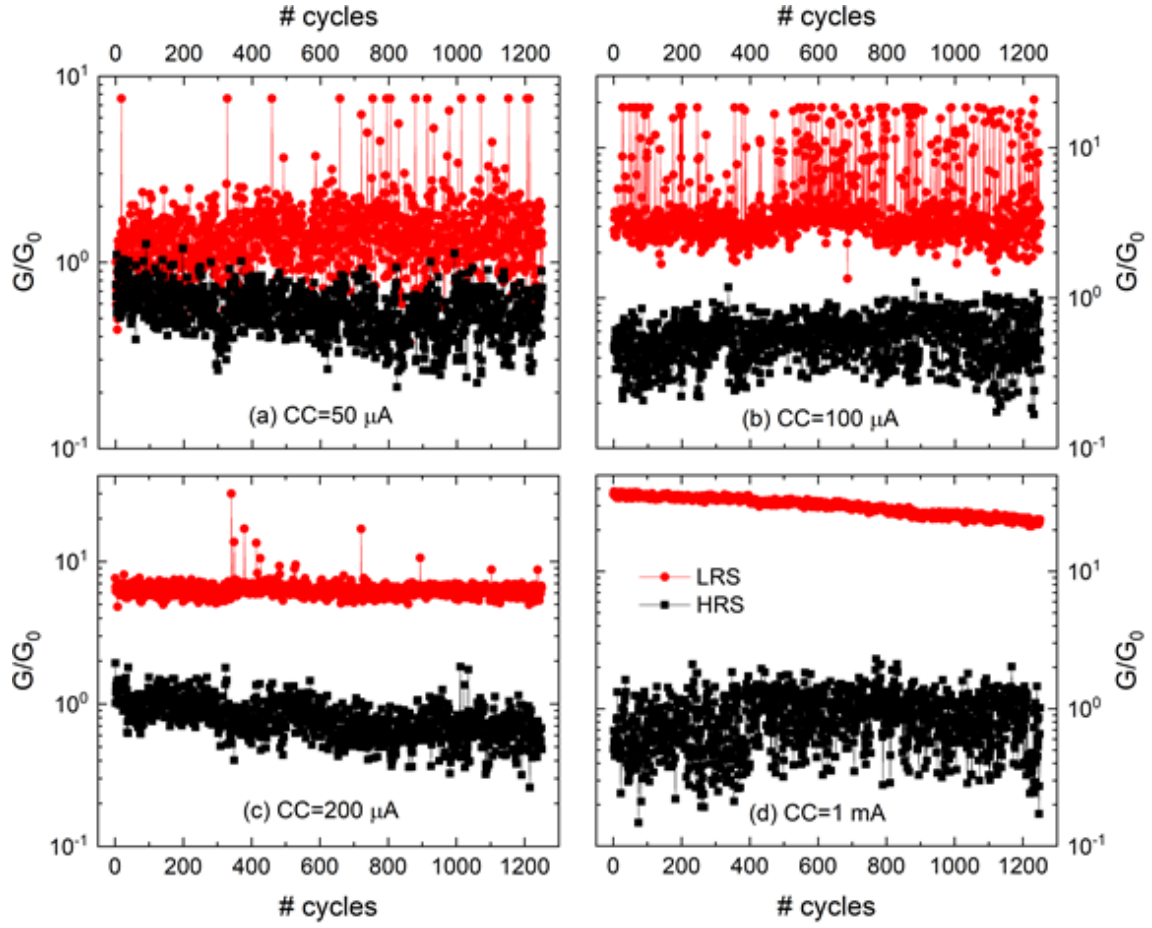

Figure 4.3 Evolution of the normalized conductance (measured at an applied voltage of 0.1 V) with the number of cycles of voltage the sample was subjected to, for the four CC values used.

It is observable (Figures 4.2 and 4.3), that the quality of switching (here defined as the ratio between the LRS and the HRS) is improved with the increase of the applied compliance. For a CC of 1 mA, one obtains a ratio $G_{HRS}/G_{LRS} > 10$, an acceptable value for most practical applications of this technology. While the LRS conductance for this CC value dropped with cycling, the ratio stayed firmly above this value, and no switching failures were observed.

Instead of going for higher values (in fact, many devices use no CC, and rely on the series resistance to limit the current through the device), lower values of compliance were used to better observe the transition from the HRS to the LRS.

The frequency of overshooting (the current during the RESET operation going above the value of the CC) tended to decrease with the increase of the CC. The upper tail in the conductance distributions decreases with increasing CC (Figure 4.2). This can also be observed in the evolution of the SET and RESET voltages (Figure 4.4). There, the SET voltage can be seen to decrease with increasing CC. It was found that generally, higher SET voltages led to overshooting cycles, which can be related to a more abrupt forming of the filament. This tends to be associated with the discharge of the parallel parasitic capacitances, from sources like the connecting cables [111–114].

The RESET voltage (defined as the voltage where the maximum current value occurs) also shows an evolution towards lower voltages with increasing CC. It is particularly visible in the 200 μA and 1 mA curves, and also in the respective graphics in Figure 4.1, that there appears to be two RESET events, one at roughly 0.5 V and one at roughly 0.8 V. This could be evidence of the presence of multiple conducting filaments [115], but no other indications like a stepwise SET transition [116] can be observed. It is also possible that other competing processes are manifesting, like the reported [117] exchange of oxygen with the AE.

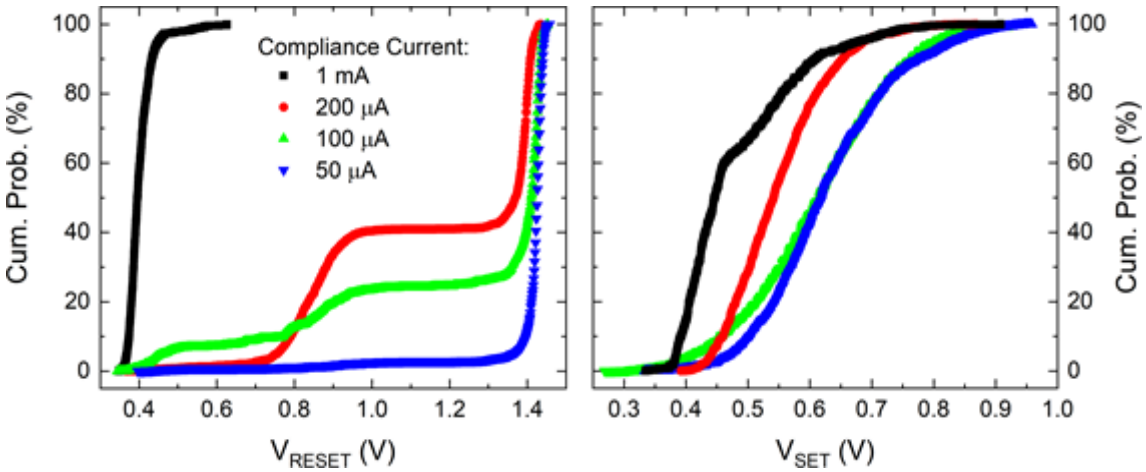

Figure 4.4 Cumulative probability of the RESET (left) and SET (right) voltages, for the four CC values used. The SET voltage was here defined as the voltage for which there was a jump of conductance of at least an order of magnitude, or, if such a case did not occur (especially for lower values of CC), the value when the current reached the CC. The RESET voltage was here defined as the voltage for which the current reached its maximum absolute value.

The peak at 1.5 V, which grows with decreasing CC, is the result of the RESET transition being less abrupt. For the lower CC values, especially for 50 μA, the two resistance states are so close (even overlapping at times – Figures 4.2 and 4.3) that pinning down where the transition occurs becomes increasingly difficult, and even raises into question whether a transition is taking place at all. This criterion for RESET voltage was used, even with this limitation, as it works fine for the higher values of CC, and the cycles where it does not are those where, due to said overlapping of the LRS and HRS, looking for it is not interesting.

4.2 Convergence and guess influence in the QPC model's application

For studying the QPC model, the single subband approximation expressed in (2.10) was used. In all QPC simulations presented throughout this thesis, a value of $\beta = 1$ was used. As per the VCM model previously introduced, this depicts a constriction at the disk, close to the anode, where the migration of vacancies widens or thins out the bottleneck. An advantage of this assumption is reducing the number of variables present in the model, so as to ease in solving the numerical problem and minimize overfitting. The choice of $\beta = 1$ over $\beta = 0$ was due to the polarity of the SET and RESET operations, which indicate that the AE is the BE. The position of the scavenging Ta layer at the TE also corroborates this assumption, as this layer, which is the TE, should constitute the CE.

Starting with an initial overlook of the problem: after a direct application of Newton's method, one can compare the norms of the first ($n1$) and second order ($n2$) components of the Hessian matrix (Table 4.1). These are defined as the left and right members of the expression (3.11); the norms were defined as the $p-2$ norms of said matrices (the $p-1$ and $p-\infty$ norms were also tested, and returned similar results). The calculations were performed using initial guesses covering 18 possible combinations of values, spanning a wide range for each parameter. Both the original QPC model and the single subband model were tested this way (as the fixed barrier height is fundamentally a simplification of the single subband model, merely removing a variable, this model was not tested). In fitting the original model, the number of subbands was limited, up to the point when adding another one had a contribution of less than 0.1 to the calculated current.

It is observable that the small residuals condition is not observable, and in some cases the second order component has a greater weight than the linear one. As such, the use of the Gauss-Newton method alone is not recommended, because without the second order contribution convergence may not be achieved.

$(N/V_0, \Phi, \alpha)$	**(1/2, 2, 1)**	**(1/2, 1, 1)**	**(1/2, 0.5, 1)**	**(1/2, 2, 3)**	**(1/2, 1, 3)**	**(1/2, 0.5, 3)**
n1/n2 original model	1.047	1.567	3.009	0.8881	1.061	2.210
n1/n2 single subband	1.459	1.459	1.837	1.149	1.521	1.892
$(N/V_0, \Phi, \alpha)$	**(1/2, 2, 5)**	**(1/2, 1, 5)**	**(1/2, 0.5, 5)**	**(10/0.5,2,1)**	**(10/0.5,1,1)**	**(10/0.5,0.5,1)**
n1/n2 original model	0.9277	0.9723	1.386	1.053	2.214	7.008
n1/n2 single subband	1.040	1.740	2.144	1.712	0.6894	0.5098
$(N/V_0, \Phi, \alpha)$	**(10/0.5,2,3)**	**(10/0.5,1,3)**	**(10/0.5,0.5,3)**	**(10/0.5,2,5)**	**(10/0.5,1,5)**	**(10/0.5,0.5,5)**
n1/n2 original model	0.9901	9.107	28.71	0.9887	16.00	9.910
n1/n2 single subband	1.136	1.975	3.956	1.039	2.340	10.87

Table 4.1 Values of $n1/n2$ obtained for 18 sets of starting $\left(N/V_0\ (\mathrm{V}), \Phi\ (\mathrm{eV}), \alpha\ (\mathrm{eV}^{-1})\right)$ guesses, for both the original QPC model (where V_0 (V) was used) and the first subband only approximation (where N was used). In the original model, a value of ω_y was used so that the height of the first subband barrier, $\hbar\omega_y \cdot 3/2$, equals the displayed value of Φ.

However, certain methods can be applied to correct the Gauss-Newton method without requiring the full Hessian calculation [104,105]. In this thesis, a simple application of the Levenberg-Marquardt method was used for this purpose. The steepest descent correction aids in selecting the searching direction, and saves processing time, especially in the case of the original model where each calculation in the Hessian must be computed for each considered subband.

To test the Levenberg-Marquardt method versus the full calculation of the Hessian, one must first devise a way to determine the value of the weight parameter λ. An additional hurdle is that the Jacobian matrix changes significantly from cycle to cycle, and an appropriate weight for one may be too large, or too small, for another. A heuristic method was thus designed to make λ dependent on $\boldsymbol{J^T \cdot J}$, so that

$$\lambda = \lambda^* \cdot \min\left(\mathrm{diag}\left(\boldsymbol{J^T \cdot J}\right)\right). \tag{4.1}$$

This ensures that $\lambda \to 0$ when one approaches the stationary point; that it will have the greater impact around the variable that shows the slowest descent, while leaving the remaining variables relatively untouched; and that one can just define λ^* for all cycles, and λ will automatically be updated to the Jacobian matrix's scale.

Starting with $\lambda^* = 1$, both the full Hessian calculation and Levenberg-Marquardt methods were fitted for the set of 18 starting guesses previously introduced. This was performed for a single random curve, and for both QPC models. It was found, surprisingly, that the full Hessian calculation did not result in convergence for all starting parameters. Note that none of these produced a curve that was particularly close to the experimental data. For the Levenberg-Marquardt model, all managed to reach a solution. This indicates that the damping of the Jacobian has a bigger impact than the addition of the second order derivatives, and manages to provide some much needed robustness to the model. As such, the Levenberg-Marquardt model was applied over a calculation of the second order terms.

This procedure was then repeated for another 29 random curves, taken from both the LRS and HRS for the four values of CC used, and the reported results were found to be representative of this larger sample. Then, a comparison of the resulting parameter values between the different initial guesses was done. Taking the results for the first curve mentioned, the guesses $(0.5, 0.5, 3)$ and $(1, 2, 1)$ were selected for the original and single subband models respectively. As these are the best initial guesses for the provided curve, they were used as the reference for comparison. The largest relative difference, defined as $\delta = |sol/\mathrm{ref} - 1|$, was calculated for both models, with the Levenberg-Marquardt method, using the same regularization weight of $\lambda^* = 1$ (Table 4.2).

$(\mathbf{N/V_0}, \boldsymbol{\Phi}, \boldsymbol{\alpha})$	**(1/2, 2, 1)**	**(1/2, 1, 1)**	**(1/2, 0.5, 1)**	**(1/2, 2, 3)**	**(1/2, 1, 3)**	**(1/2, 0.5, 3)**
Original model (%)	---	---	---	---	---	---
Single subband (%)	---	0.23	0.33	0.52	0.33	0.47
$(\mathbf{N/V_0}, \boldsymbol{\Phi}, \boldsymbol{\alpha})$	**(1/2, 2, 5)**	**(1/2, 1, 5)**	**(1/2, 0.5, 5)**	**(10/0.5, 2, 1)**	**(10/0.5, 1, 1)**	**(10/0.5,0.5,1)**
Original model (%)	---	---	---	---	6.8	1.2
Single subband (%)	220	0.18	0.45	67	67	67
$(\mathbf{N/V_0}, \boldsymbol{\Phi}, \boldsymbol{\alpha})$	**(10/0.5,2,3)**	**(10/0.5,1, 3)**	**(10/0.5,0.5,3)**	**(10/0.5,2,5)**	**(10/0.5, 1, 5)**	**(10/0.5,0.5,5)**
Original model (%)	---	6.8	---	---	---	---
Single subband (%)	68	68	66	1.4	68	61

Table 4.2 Values of $\delta = |sol/\mathrm{ref} - 1| \cdot 100\%$ obtained for 18 sets of starting $\left(N/V_0\ (\mathrm{V}), \Phi\ (\mathrm{eV}), \alpha\ (\mathrm{eV}^{-1})\right)$ guesses, for both the original QPC model and the first subband only approximation. The $(0.5, 0.5, 3)$ and $(1, 2, 1)$ guesses were used as a reference for the original and single subband models, respectively. The dashed boxes indicate cases where convergence was not achieved.

The single subband model managed to achieve a solution for all starting guesses considered. The δ values under 1% converged to the same solution of $(N, \Phi, \alpha) = (7.519, 1.468, 2.055)$.

The values under 100% converged to a different local minimum of $(12.49, 1.871, 1.887)$. The last case found a different solution in $(23.76, 2.351, 1.752)$. In all cases, it was found that $R^2 > 0.995$, indicating a very high goodness of fit. A solution different from the initial one was only found by changing the parameter N (the only outlier corresponds to a guess of $(\Phi\,, \alpha) = (2, 5)$, where the barrier is tall and extremely wide). This parameter is more loosely correlated to the solution: in fact, looking at (2.10), one can see that dI/dN is not a function of N, and $d^2I/d^2N = 0$. Upon testing further with different values of N, it was found that this parameter has the highest influence in the solution, but with a starting guess of $N < 4$ the same solution was still reached.

The original model failed to converge an alarming number of times. The original solution of $(V_0, \Phi\,, \alpha\,) = (0.4820, 0.2295, 2.270)$ was only reached once more, and the cases with $\delta > 6\%$ managed to reach a slightly different case of $(0.4820, 0.2295, 2.270)$. Further testing with different guesses revealed that the solution was almost entirely dependent on the value of V_0 (much like N, this is the fitting parameter in this model), but for convergence to be achieved a very good initial guess was required. The goodness of fit was also slightly lower, with some values being close to $R^2 \sim 0.99$. Also, for the LRS with linear curves, this model never succeeded in adequately modelling the experimental data.

Both cases were then tested with other curves, and the results reported here were found to be representative of the larger sample. Overall, the original model's implementation suffers from a lack of robustness, and a strong dependence of the solution on the initial guess. For N, one can also default to the $N = 1$ case as the initial guess, given that the model is meant to represent a single subband. For V_0 on the other hand, is harder to determine what initial guess to use, a problem compounded by small variations in this parameter having a sizeable impact on the fitting procedure. This makes deciding what a good initial guess is even harder.

Different values of λ^* were then tested in both cases. For the single subband case, it was found that values in the range of $\lambda^* \in [0.1, 10]$ provided the same solutions as the ones previously shown, with a maximum deviation of $\sim 10\%$. Smaller values converged to the no regularization case, where convergence was often not reached; higher values tended to heavily distort the solution. The LOOCV also returned good results, with $R^2 > 0.995$. This can be justified by an analysis of the $\boldsymbol{J^T \cdot J}$ matrix: looking at its diagonal, the value corresponding to N tended to be 1 to 4 orders of magnitude smaller than those matching the other two variables, which were usually less than 1 order of magnitude apart. This ties to the initial observations of a lack of correlation between N and the other two parameters. The correction to this value thus barely affects the other two, justifying why N is the main influencer of the solution.

On the other hand, almost any range of λ^* chosen failed to significantly improve the convergence for the original model. Looking at the same diagonal of $\boldsymbol{J^T \cdot J}$, one observes that, this time, all values were within 1 order of magnitude of each other. This relates to the fact that, in equation (2.9), all three of the variables are within the same exponential, and translate very similar physical quantities.

An adequate way to apply the original model was not found. Thus, only the single subband, and its derivation with a fixed barrier height, will be applied throughout this thesis.

4.3 Comparison between the QPC and conventional models

The conduction models previously introduced in Chapter 1, and considered for conduction in the type of samples used, can be grouped as per the way the current depends on the applied voltage (assuming a constant temperature), so that a linearized form of the expressions can be used for the fit (Table 4.3).

The direct tunnelling and ionic conduction mechanisms were discarded. Considering the oxide thickness and maximum voltage used, the Fowler-Nordheim tunnelling mechanism was deemed to be the most likely tunnelling scenario, without the influence of traps. Given that the direct tunnelling expression considered could not be easily linearized, only the Fowler-Nordheim and trap-assisted tunnelling scenarios were investigated further. Ionic conduction is unlikely, as, at room temperature, electronic mobility is greater than oxygen ionic mobility in TaO_x [118,119].

Trap-assisted Tunnelling (TAT)	$\ln(I) = -A\frac{1}{V} + B$
Fowler-Nordheim (F-N) Tunnelling	$\ln\left(\frac{I}{V^2}\right) = -A\frac{1}{V} + B$
Schottky Emission	$\ln(I) = A\sqrt{V} + B$
$I \propto V^p$ (Ohmic, NNH and VRH)	$\ln(I) = pV + A$
SMSE and Poole-Frenkel Emission	$\ln\left(\frac{I}{V}\right) = A\sqrt{V} + B$

Table 4.3 Linearizations performed in order to evaluate goodness of fit of the models presented in Table 1.2. I is the electric current, V the voltage, and A and B constants.

The SCLC, which is not linearizable on account of the threshold change in power, was evaluated facing the experimental data for the HRS (Figure 4.5). Given that no abrupt transitions were observed, it is then natural to assume that, if the SCLC model is to hold, the nonlinearity observed must result from the transition between the linear regime and the traps-unfilled SCLC regime, as introduced in Chapter 1. As such, a sample of 30 curves was taken over the four used compliance values, and found to be representative of the source population. The curves were then partitioned into two segments. Both were then fitted individually via least squares estimation as per $I \propto V^p \Rightarrow \ln(I) = p\ln(V) + C$. As per the SCLC model, the first segment is expected to result in $p_1 \sim 1$, and the second in $p_2 \sim 2$. As such, the residual ${r_p}^2 = (p_1 - 1)^2 + (p_2 - 2)^2$ was then calculated for all possible partitions where each segment had at least 3 points (to avoid the trivial 2 points curve scenario). The best "SCLC fit" was then the one that minimized ${r_p}^2$. The two drawn curves in Figure 4.5 are representative of the results found, for more linear curves, and more nonlinear curves. The results show that the SCLC is not a good fit for the experimental data: in the more linear case, it is found that the data can be fitted rather appropriately by a single region with $p = 1.14$, and a second region is only obtained due to noise in the data. As for the more nonlinear curve, while acceptable values for p_1 and p_2 were found, the change in slope is only an expression of the inherent nonlinearity, and no evidence of a threshold change in slope was observable.

For the linearizations in Table 4.3, all cycles for both the LRS and HRS were evaluated, and, to measure the goodness of fit, the coefficient of determination R^2 was obtained (Figure 4.9). This measure of the residuals is here defined as

$$R^2 = 1 - \frac{RSS}{SYY}, \tag{4.2}$$

where RSS is the residual sum of squares and SYY is the data variance [120]. After applying the linearizations present in Table 4.3 and obtaining the fitting parameters, the curves were redrawn on the linear I-V basis, and it was in this basis that the value of R^2 was calculated. This allows for a better comparison between the different mechanisms, as the coefficient of determination is not conserved for nonlinear transformations (these do not necessarily conserve the ratio between the data variance and the fit residuals).

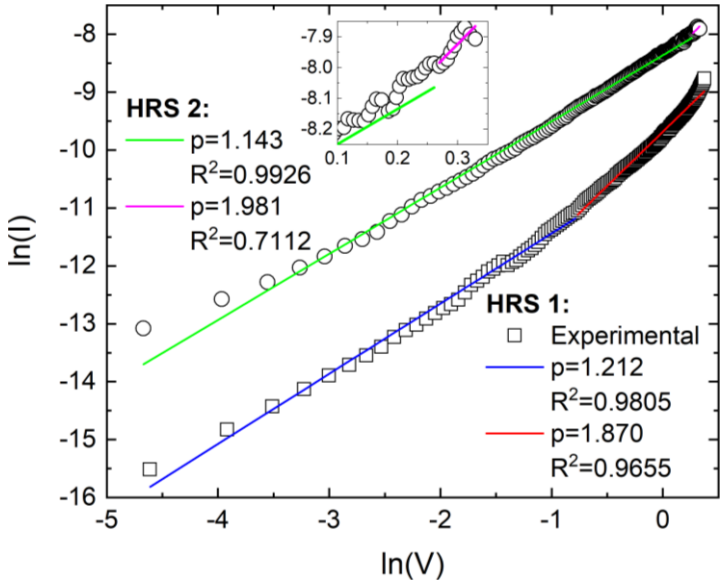

Figure 4.5 Results of the SCLC fit for a more nonlinear (HRS 1) and linear (HRS 2) HRS curves. The curves were partitioned into two segments, fitting them to $I \propto V^p$ and calculating the resulting powers p_1 and p_2. A residual ${r_p}^2 = (p_1 - 1)^2 + (p_2 - 2)^2$ was minimized to find the best partition. The inset figure highlights the second segment for HRS 2.

The values of R^2 found were relatively low (with $1 - R^2$ not going below 10^{-5}). Also, all models featured a similar amount of parameters (2, except for the QPC which has 3), which was also small when compared to the number of data points available (150 for the HRS, and between 30 and 100 for the LRS). As such, it was determined that more extensive criterions for goodness of fit in nonlinear fitting, like the Akike and Bayesian criterions [120,121], were unnecessary. Applying corrections to the coefficient of determination was also considered not relevant, as the simple R^2 is easy to calculate, widely used and very familiar, and allows for the necessary discrimination.

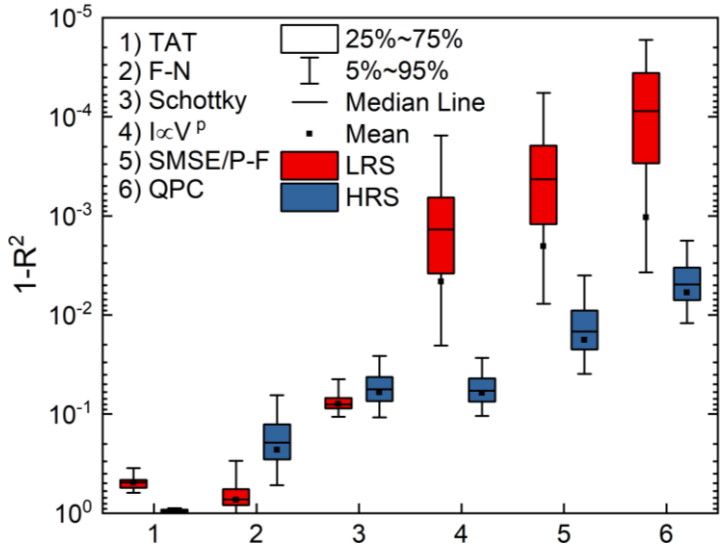

Figure 4.6 Distribution of the indicators of goodness of fit obtained, expressed as $1 - R^2$. These were taken for both device states, for the various considered models, over all cycles taken. The QPC model used was the single subband one.

After comparing the models' quality of fit, the TAT and F-N were found to be poor fitting models, on account of the very low values of R^2 found. Only the other 4 were then evaluated more carefully. To do so, the sample of 30 curves used in the SCLC evaluation was taken, fitted to the 4 models, and each was individually inspected. In Figures 4.7 and 4.8, a pair of these curves for each resistance state, one of the curves being more nonlinear, and the other being more linear, are displayed. The findings are representative of those encountered in the whole sample.

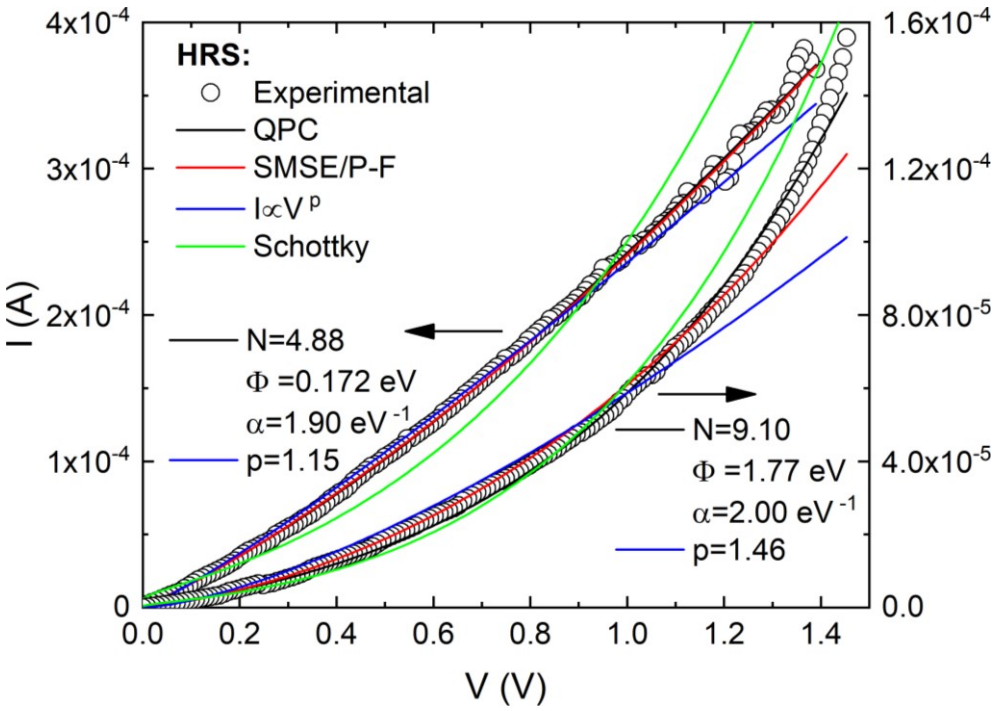

Figure 4.7 $I - V$ characteristics obtained in the HRS, for more nonlinear (right) and more linear (left) curves. These were found to be representative of the rest of the data. The lines show the results of the QPC, SMSE/P-F, $I \propto V^p$ and Schottky fits. The values of N, α and Φ are those resultant from the QPC fit, and the value of p is the power of the $I \propto V^p$ fit.

In the HRS (Figure 4.7), it is observable that the Schottky emission is not a very appropriate fit for both the nonlinear and linear cases: it proves unable to accurately represent the curvature in both the linear and the nonlinear scenarios. The $I \propto V^p$ is a better fit for the more linear data, with the value of $p = 1.15$ being close to the expected value of 1. However, it fails to properly account for the nonlinear curve, and the resulting value of $p = 1.46$ is rather far from the expected. In this case, both the QPC and SMSE/P-F are fairly good fits for both curves, with the QPC being able to reproduce the nonlinear data slightly better.

In the LRS (Figure 4.8), it is again observable that the Schottky emission is not an appropriate fit in both scenarios. The $I \propto V^p$ group of models is now a better candidate, being able to fit the linear data very well with the value of $p = 1.01$ being the one expected; as for the nonlinear data, a value of $p = 1.30$ is obtained, and while the goodness of fit has improved from the HRS case, it is still not an adequate model to describe this data. Again, both the QPC and SMSE/P-F are great fits for the presented data, and this time are close to indistinguishable.

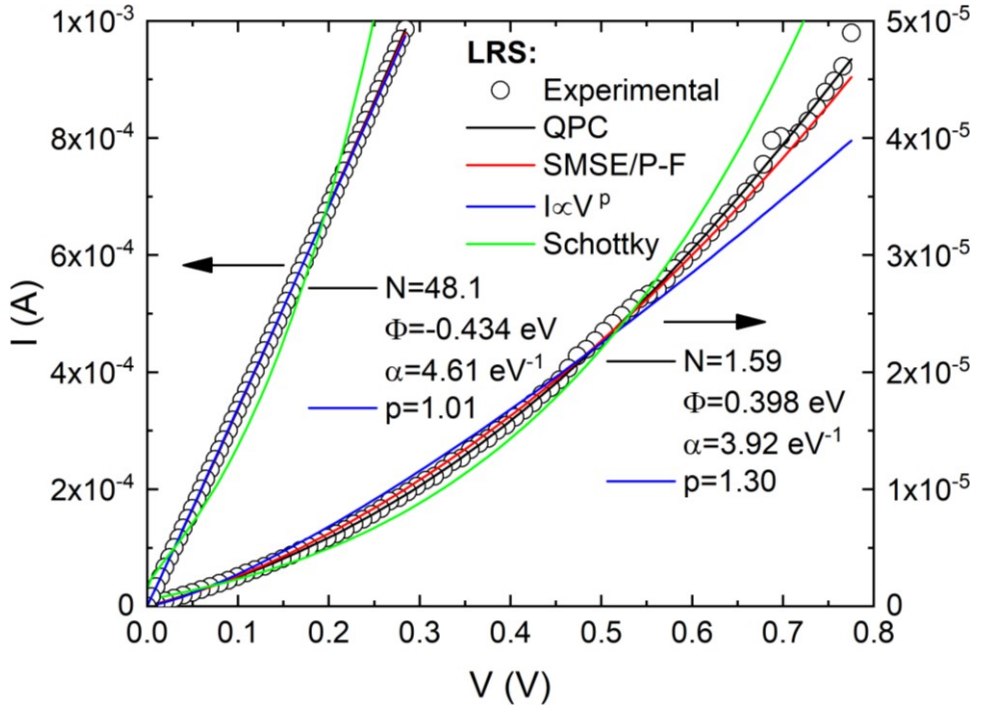

Figure 4.8 $I-V$ characteristics obtained in the LRS, for more nonlinear (right) and more linear (left) curves. These were found to be representative of the rest of the data. The lines show the results of the QPC, SMSE/P-F, $I \propto V^p$ and Schottky fits. The values of N, α and Φ are those resultant from the QPC fit, and the value of p is the power of the $I \propto V^p$ fit.

A closer look at the SMSE/P-F models results in

$$A = \frac{1}{r}\frac{e}{kT}\sqrt{\frac{e}{d\pi\varepsilon_0\varepsilon_r}}, \tag{4.3}$$

where A is the parameter obtained via linearization included in Table 4.3, e the electron charge, k the Boltzmann constant, T the temperature, d the thickness of the active oxide (over which it is assumed the entire voltage V is applied, so that the field E is given by $E = V/d$), ε_0 and ε_r the vacuum and relative permittivities respectively, and r is a compensation parameter ranging from 1 (normal P-F emission) to 2 (P-F emission with compensation) [122,123]. One shall consider the simple case that $d = 5\,\text{nm}$, meaning that the voltage is applied over the entire oxide layer; according to the VCM model, it would be applied mostly in the disk region, resulting in a much lower value of d. One can then extract the expected value for the relative permittivity, accounting for the possible compensation, via

$$r\sqrt{\varepsilon_r} = \frac{1}{A}\frac{e}{kT}\sqrt{\frac{e}{d\pi\varepsilon_0}}. \tag{4.4}$$

The results for the LRS and HRS, considering only those above the 5% line for R^2 displayed in Figure 4.6 were then obtained (Figure 4.9).This measure was taken to avoid outliers resultant from effects like noisy data, where the fit was not appropriate in the first place. Even accounting for a maximum value of $r = 2$, the results obtained are far larger than $\varepsilon_r \sim 20$, as reported in the literature for amorphous [123–129] tantalum oxide thin films, and even up to $\varepsilon_r \sim 40$, as reported for non-amorphous [130–134] thin films. The results would also be much larger if one was to

lower the value of d, according to the VCM model, as previously discussed. These results also display a very high variance (especially in the LRS), and sometimes display negative values, with a very good quality of fit (again, particularly in the LRS). This indicates that, while the mathematical equation can provide a good fit for the data, the SMSE or P-F emission models may not be adequate to describe the physical phenomenon present in the device, even in the HRS where this mechanism would be more likely.

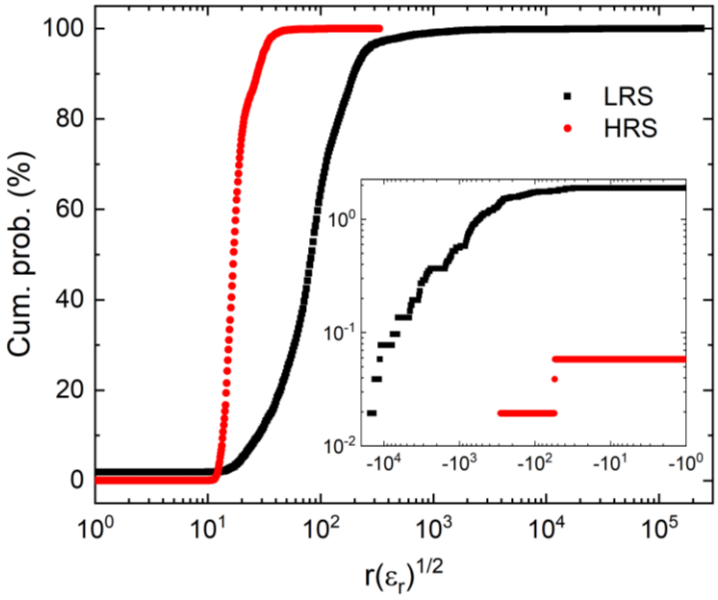

Figure 4.9 Cumulative probability of the values of $r\sqrt{\varepsilon_r}$ obtained from the SMSE/P-F fits of the experimental data. These were taken for both device states, in accordance to equation (4.4) using $d = 5$ nm.

On the other hand, the values of $\left(N, \Phi\ (\text{eV}), \alpha\ (\text{eV}^{-1})\right)$ resultant from the application of the QPC, (Figures 4.7 and 4.8) were: $(9.10, 1.77, 2.00)$ and $(4.88, 0.172, 1.90)$ for the HRS, and $(1.59, 0.398, 3.92)$ and $(48.1, -0.434, 4.61)$ for the LRS. The first trio of values for each resistance state correspond to the more nonlinear curve, and the second to the more linear one. The first three sets of values are consistent with the considerations on the model presented so far, and the observed lowering of the barrier height (Φ is decreasing) between the HRS and LRS, and between the nonlinear and linear curves, all match the earlier predictions. The fourth one, for the linear LRS curve, ties into the observations done in the previous section, where the model converges into linear conduction (high N and α, low Φ).

One can thus conclude that the QPC appears to be the only analysed model capable of fitting both the LRS and HRS, with an elevated goodness of fit, and possibly providing a proper phenomenological explanation for the transition between both states.

4.4 Temperature measurements and observed RTS

In the QPC model's introduction in Chapter 2, it was made an assumption that there is no smearing of the Fermi function at the electrodes. Upon simulating the distribution for different temperature values (Figure 4.10, left), it was found that at 300 K there is already some spreading as compared to lower temperatures. However, the difference is rather small, and only at very high temperature values (> 1000 K) does this effect become particularly noticeable.

The current I was also simulated without applying the zero-temperature limit, according to equation (2.6), versus the temperature. The relative difference of its value at a temperature T to the value at 1 K, $I_{1\,K}$, given by $\delta_I = |I/I_{1\,K} - 1|$, was then calculated (Figure 4.10, right). An analysis of the results shows that the effect of the smearing is rather small,: at $300\,K$, this

difference was always smaller than 1%. This holds true even for a high barrier height of 2 eV: the bandgap of Ta_2O_5 is ~4 eV [99,126,135], so 2 eV is about the distance between the Fermi level and the bottom of the conduction band. The values of α and Φ used cover both the LRS and the HRS, and are consistent with the results of the model's application found in later sections.

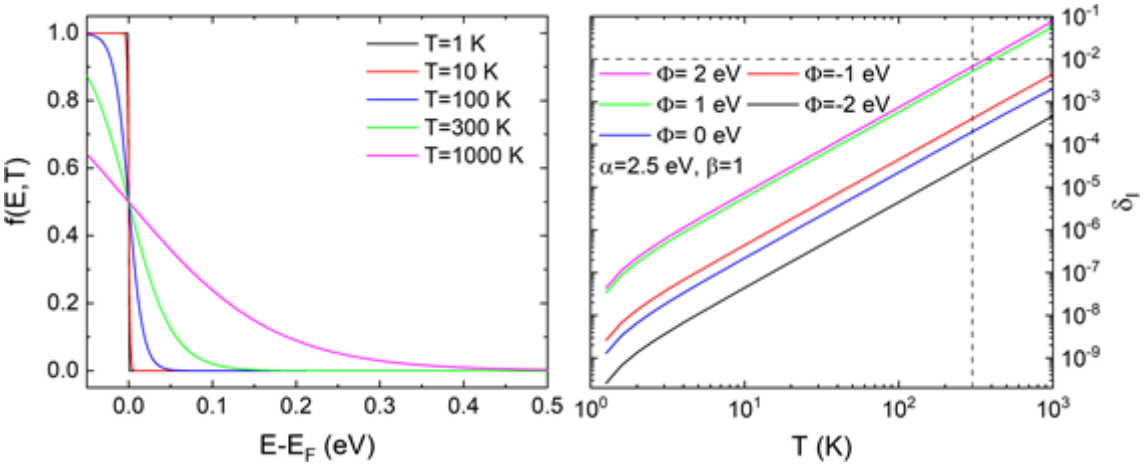

Figure 4.10 (Left) Evolution of the Fermi-Dirac distribution for different temperatures, versus the energy measured in relation to the Fermi level; (Right) Simulation of the relative difference of the current from its value at 1 K, at 0.1 V, defined as $\delta_I = |I/I_{1\,K} - 1|$. The current was simulated as per equation (2.6), with $N = 1$ and the other parameters taking the values displayed, at an applied voltage of 0.1 V.

The original model introduced in Chapter 2 does not predict another direct dependency of the current on temperature. One can consider that the thermal excitation of charge carriers results in an effective lowering of the potential barrier [136]. This effect should be particularly noticeable in the HRS. In the LRS, the influence of the series resistance's temperature dependence should be more noticeable. As such, one should observe an increasing conductance with temperature increase in the HRS, and the metallic behaviour of the plug in the LRS.

To study the evolution of the HRS with temperature, a small triangular voltage ramp was applied to the sample, only up to $0.1\ V$ in order to prevent further resetting. The voltage applied had a negative value, so as to not induce a change to the LRS state. The temperature was first lowered to $10\ K$ and then raised in increments up to room temperature. However, abrupt transitions in the current for a fixed applied voltage were observed, independently of the temperature at the moment. To further explore this phenomenon, $140\ I-V$ sweeps up to $0.5\ V$ were measured, at a fixed temperature of 120K. The increase in stopping voltage was found not to significantly exacerbate the problem. This led to the assumption that the issue at hand was possibly due to the power supply, and interference with the measurement setup, instead of being related to the temperature.

The resulting values of conductance at the stopping voltage of $0.5\ V$ (Figure 4.11) clearly display this phenomenon. The conductance appears to switch between five levels at roughly $G/G_0 = 1.15, 1.10, 1.08, 1.035$ and 1.00, in an overall descending fashion. The intermediate values encountered after the 100^{th} cycle are a result of smaller fluctuations near the end of the curve, and not representative of the overall curve, unlike the five distinct levels indicated (Figure 4.12).

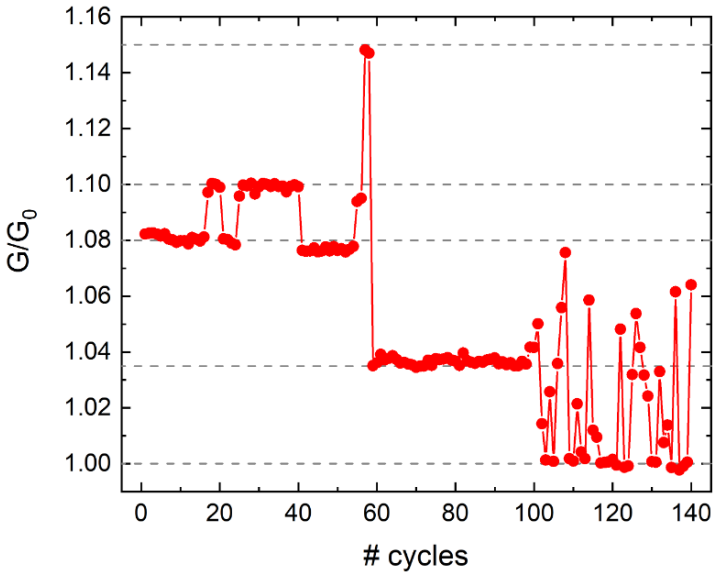

Figure 4.11 Evolution of the sample conductance with cycle number, displaying its abrupt nature. The five levels at $G/G_0 =$ 1.15, 1.10, 1.08, 1.035 and 1.00 are highlighted with dashed grey lines.

When the curves were superimposed, the five levels observable in Figure 4.11 were determined and fitted by the QPC model (Figure 4.12). The starting guess used in all five cases was $\left(N, \Phi\ (\text{eV}), \alpha\ (\text{eV}^{-1})\right) = (1,2,1)$. The results displayed a very slight change in α from the lowest conductance case (black line) to the highest (violet line), of 1.653 to 1.911. This change does not produce a significant change in the transmission probability. The value of N displayed a higher change, rising from 3.551 to 4.068. The value of Φ displayed the highest variation, lowering from 1.273 to 0.7487. Overall, α appears to not have changed significantly, while N decreases and Φ increases in a more significant fashion. Going back to the full QPC model, this can be interpreted as the energy separating the multiple subbands being lowered (the constriction widens), with more of them contributing to conduction as a result. This means the fitting parameter V_0 would remain rather constant, while the parabolic well curvature ω_y would decrease. The interpretation of the model's parameters is further explored in Sections 4.4 and 4.5.

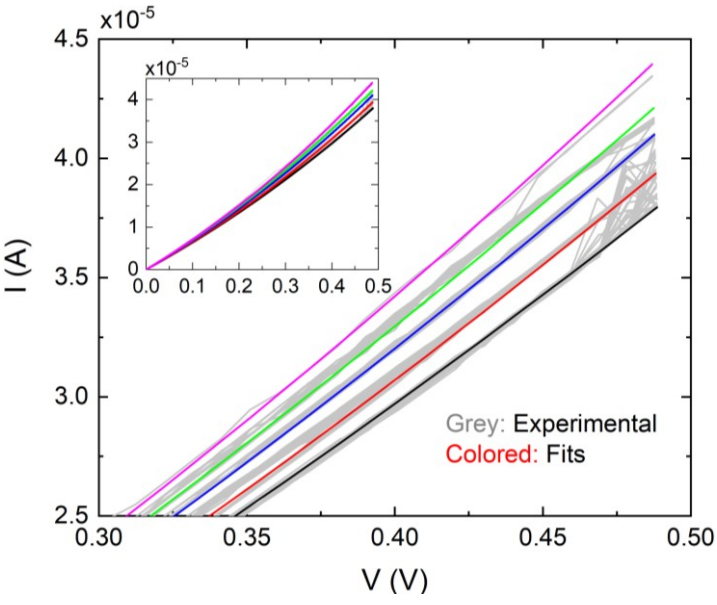

Figure 4.12 Experimental data resultant from the 140 curves obtained at $T = 120\ K$. The inset figure contains the overall curves, over the entire range of the applied voltage. The five regions previously introduced were fitted via the single subband QPC approximation, and the resulting curves displayed in colour. The mentioned transitions around $0.5\ V$ are clearly displayed, but multiple transitions between the regions were found along the entire range of applied voltage values.

These transitions display similarities with the random telegraph signal (RTS) phenomenon [137–139], commonly observed in oxide thin film. The RTS usually relates to the trapping and de-trapping of electrons, providing two levels for the current where the carrier is trapped (lower

I) or emitted (higher I). With a continuous flow of current through the device, as was the case with the multiple applied smaller cycles, sources like random thermal excitations or noise in the power source can cause the jumps between these allotted states. One can extrapolate that, if one was to consider multiple traps being present in the material, more than two levels could be achieved, as is observed in Figure 4.12. Given that, in a VCM device such as the one tested, oxide defects play a fundamental role in the conduction, the existence of these multiple traps is to be expected.

Other interpretations could be drawn as to the origin of the phenomenon, and to the apparent evolution of the conductance (a truly random process would not tend to lower values). It is also possible that, with the continuous application of negatively biased voltage, small RESET events are being induced, and the transitions observed resultant from the vacancy motion removing available traps for conduction.

Another interesting interpretation provided by Degraeve *et. al.* [79] ties this phenomenon to the evolution of the QPC model's parameters. Here, the constriction is considered to contain an integer number of particles, and the discrete states observed are resultant from random atomic motions widening or thinning the constriction. The constriction's width ties directly to the lateral quantization of energy, and to the parameter ω_y of the model. One can then pose a scenario where the successive RESET events result in a quantized reduction in the CF's width. This proposal is consistent with the earlier observations on the evolution of ω_y, inferred from the observed evolution of N and Φ. This offers another argument for the application of the QPC model to these samples.

On top of the measurements mentioned in the previous Chapter, an APC Back-UPS Pro 550 was then added in order to ensure that the power supply to the computer and to the Sourcemeter were stable. This equipment, however, did not provide galvanic protection. After additional tests, the issue persisted, and appeared independent of the temperature at which the measurements were taken. It is thus believed that the lack of galvanic protection present in the circuitry used is leading to the flow of current through the common ground, causing small changes in the sample's resistance state when such an event occurs. This is further corroborated by the observation that the use of light switches or plugging in of electrical devices in nearby sockets can cause the sample to SET into a very low current state.

Over the course of this work, large numbers of cycles were obtained, over a focus on singling out an individual state. This has the fortunate effect of preventing such events from significantly impacting the quality of data obtained, and conclusions taken. While individual states have been shown to change abruptly when these events occur, over a large number of cycles the focus of data collection becomes to study values of statistical significance. Here, such transitions can be properly accounted for if overly abrupt, and if not, will be overshadowed by the intrinsic variance inherent to the switching phenomenon, and drowned out as noise.

4.5 Results of the single subband approximation

Looking back at the single subband model introduced in Chapter 2, one recalls that it hinges on the assumption that only the first subband contributes to conduction. This assertion was tested (Figure 4.13) by using equation (2.4) to simulate the transmission using the first five subbands (as the contribution of higher values of n is very small), for different combinations of α and Φ.

The scenario (c) stands out as one where the transmission curve displays a stepwise behaviour. Here, the transmission curves are sharp (high α), but the energetic distance between subbands is small (low Φ). As such, the stepwise fashion is precisely the result of the contribution of multiple subbands to conduction. Based on this result, it would be expected to see jumps in the measured conductance as the voltage increases, and more of said subbands were filled.

One reason this was not observed could be that we instead find ourselves in scenario (a), meaning the first subband is energetically distant enough that the following subbands are unable to contribute to conduction. This would be precisely where the single subband approximation would be successful.

Scenarios (b) and (d), where the transmission curves are smoother (lower α, also meaning a wider constriction), could also be observable. While for (b) the first subband is the main contributor to conduction, and the proposed approximation would not be too far off, one cannot claim the same for scenario (d). However, one could reach very similar curves by taking a single subband and adding an external multiplication factor to reach the $T > 1$ values required. The result would translate to multiple conduction paths, each with a single subband contributing to conduction – the same interpretation given in Chapter 2.

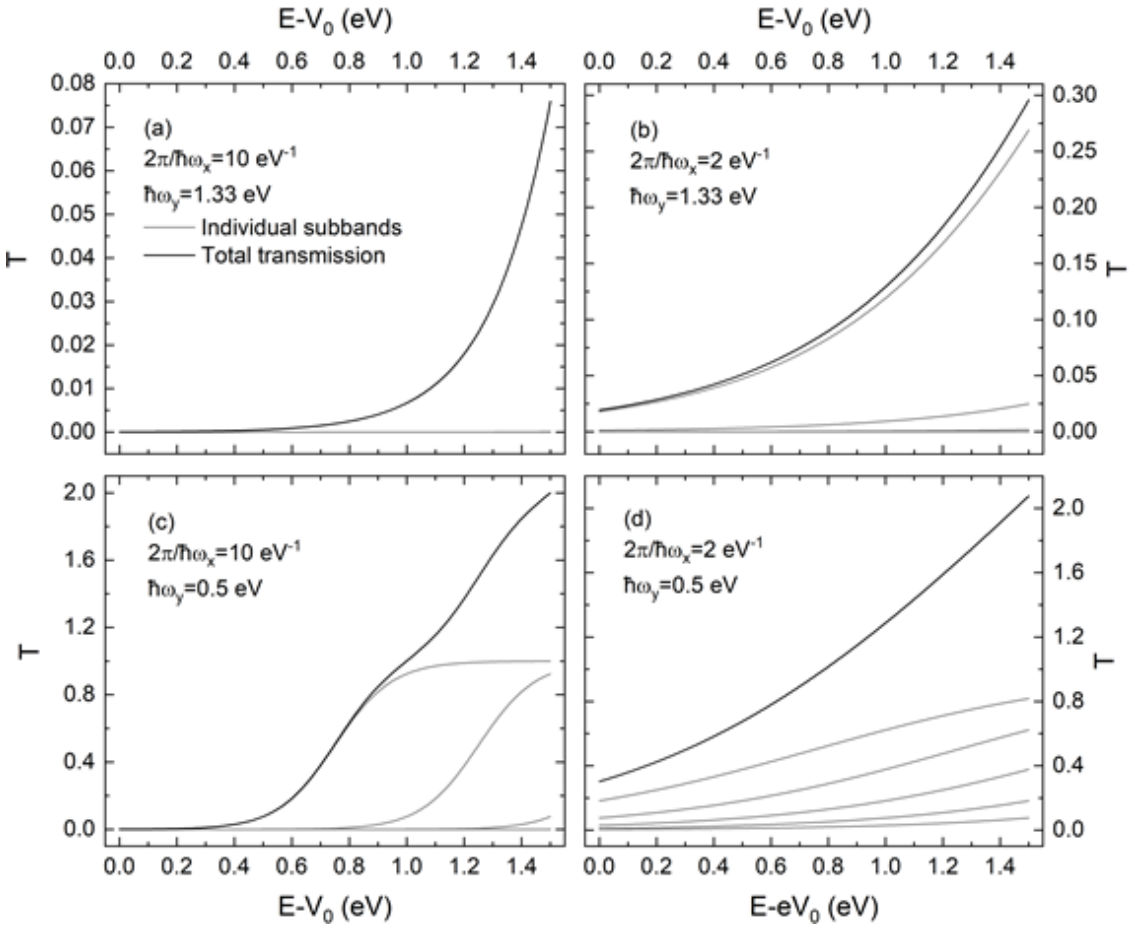

Figure 4.13 Simulation of the transmission using equation (2.4) for the first five subbands, and the resulting total transmission (the contribution of $n > 6$ is neglectable in all scenarios). Four combinations of $\alpha = 2\pi/\hbar\omega_x$ and $\hbar\omega_y$ (related to Φ) were used.

Therefore, an issue arises in interpreting the quantities obtained: N should denote the number of conductive paths, and Φ an effective measurement of their overall height. But these quantities could instead denote the effective contribution of the multiple subbands. Fundamentally, it is an

issue in distinguishing between having few parallel conductive paths with the same barrier height, and many parallel paths with decreasing barrier heights (as the multiple subbands can be thought of as parallel contributors to conduction).

This problem would be solved if the best fit occurred for $N = 1$ and a high α, as the scenario would then default to (a), where the approximation has been shown to be valid. For those curves where the best fit occurs for $N > 1$, it is recommended to look at the parameters of the single subband model, and interpret them in the light of the original multiple subband model. While the values of ω_y and V_0 can't be directly obtained from Φ and N, certain relations can be inferred. Given that N should be related to the number of subbands contributing to the conduction, it should be inversely proportional to ω_y (the more subbands contribute, the smaller their energetic distance, given by $\hbar\omega_y$, should be). As such, if both N and Φ are large, it must be that the value of V_0 is elevated. Likewise, if both N and Φ were small, V_0 would likely be very small or even negative, balancing out the expected large value of ω_y. With this kind of analysis in mind, one can still apply this approximation without losing too much physical proximity to the original model, and while still being able to comprehend the nature of the constriction.

Special care must be taken when both α and N are large. As this represents a high conductance with a wide constriction, it is likely that any curve with this pair of parameters would also display a low Φ, or the stepwise behaviour displayed in (c) should be observable. However, a large α and small Φ, as seen previously in Section 4.2, correspond to the absence of the barrier itself. One could thus be approaching the limit where the ballistic transport is no longer applicable (the electron's mean free path could become smaller than the constriction). A particular advantage of this approximation is that it can accommodate for this weakness in the physical model. In this limit, the original equation (2.9) results in

$$I \approx \frac{G_0}{e} \sum_{n=1} eV = \infty. \tag{4.5}$$

On the other hand, the expression for the first subband approximation (2.10) yields

$$I \approx NG_0V, \tag{4.6}$$

where, if $G = NG_0$, one is able to recover Ohm's Law. The mathematical model resulting from the single subband approximation thus has the advantage of being more compact. Also, one can use it even in the limit where the a quantum point contact may not exist: given that, in Ohmic conduction, $I = GV$, and, assuming a cylindrical shape to the filament with radius r and length l, one can take equation (4.4) and obtain $I \approx NG_0V = \sigma\pi r^2/l \cdot V$. Here, the resulting parameter N would instead be an indicator of the filament's width.

The single subband model was then applied to all curves (Figures 4.14 to 4.17). A single set of starting parameters, $\left(N, \Phi\ (\text{eV}), \alpha\ (eV^{-1})\right) = (1,2,1)$, was used for all curves, and the regularization weight chosen was $\lambda^* = 1$. The system was then solved using a trust-region reflective algorithm [140].

The results of N (Figure 4.14) show an apparent lack of correlation between this parameter and G for the HRS. For the LRS, it can be seen that, at higher conductance values, these values fall over the $N = G/G_0$ curve, denoting a transmission probability of $T = 1$. This finding is consistent with the physical root of the model, as a high conductance state should be achieved through the disappearance of the potential barrier. This also means that higher conductance curves tend to be more linear.

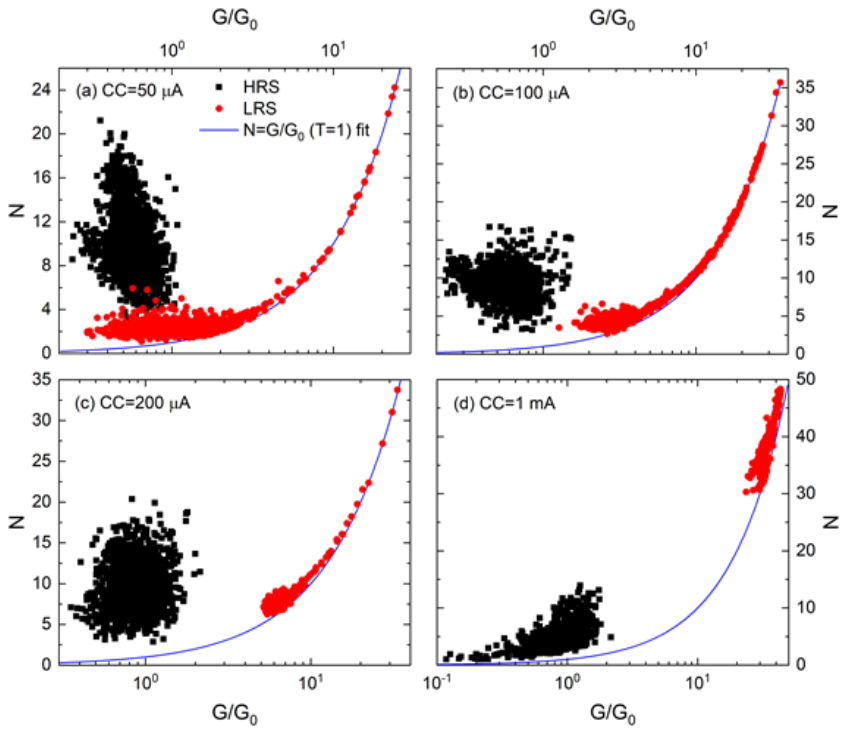

Figure 4.14 Values of N resulting from the application of the single subband QPC model. They are plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed.

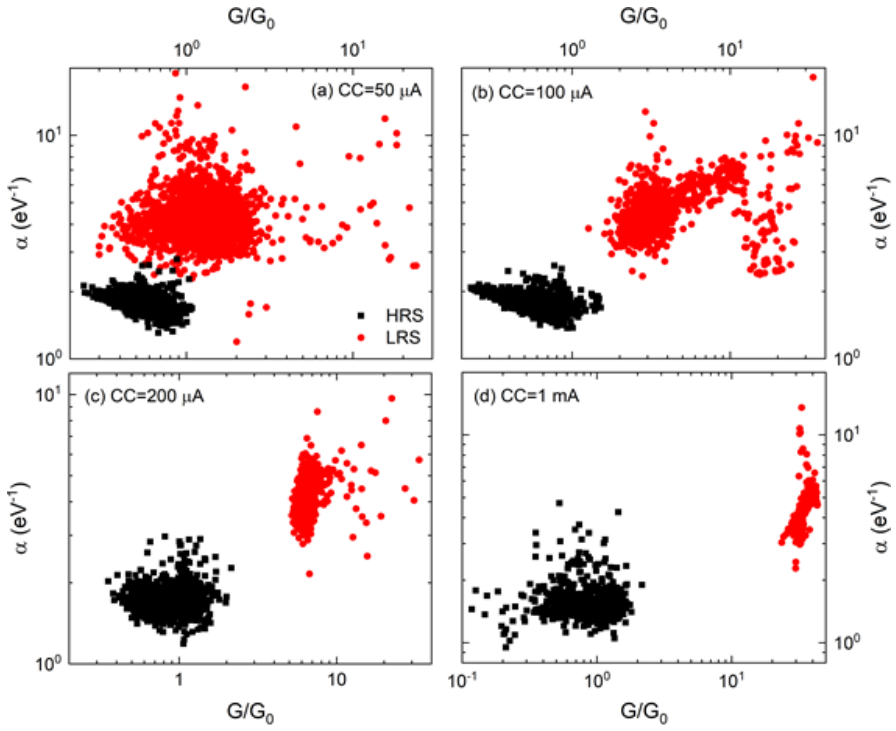

Figure 4.15 Values of α resulting from the application of the single subband QPC model. They are plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed.

The CC appears to have little impact on the HRS. The only significant change between the four cases is that, for 1 mA, a tail tending towards lower conductance values appears. On the other hand, it appears to act as a selecting mechanism for the LRS current: higher values of CC result

in overall higher conductance, and on the total disappearance of lower conductance values. Models have been presented [141–143] suggesting that this is due to the fact that higher CC values result in more damage being done to the oxide. The extra BD events result in the joining of smaller CFs into thicker units, which results in a thicker filament being formed. It could thus be that, at higher conductances, one can no longer state that the electron mean free path is larger than the constriction width, and ballistic transport no longer applies. Thus, one falls on the limit discussed in equation (4.4), where the physical basis for the QPC model is no longer applicable, but its mathematical form is still an appropriate fit to the data.

The results of α (Figure 4.15) do not show a strong correlation between this parameter and G in either state. The strongest relation between the two appears in the HRS for a CC of 50 and 100 μA, where one could attempt to define a correlation between the two, according to $\ln(\alpha) = -A \cdot \ln(G) + B$. The apparent descent, however, is very small, and not present in any of the other CCs; it may be merely an artefact of the fitting, possibly resultant from the initial guess being close to 1. Most likely, it is a tendency towards the $\alpha \to 0$ limit discussed in Section 4.2: while this may not be a sound solution from a physics point of view, it is a minimum of the mathematical problem for more linear curves. This tendency could thus be merely an indication of the increasing linearity of the curves with increasing conductance.

Here, the CC does not appear to have any effect on the value of α, apart from the high conductance LRS cycles displaying higher values. This is consistent with the previous assertion of a widening of the filament, which would result in an increase in this parameter's values. However, if one is indeed at the limit where the QPC model is no longer physically viable, this could be a mere mathematical artefact of the transmission tending to 1 due to the high linearity of the curves being fitted. This would naturally mean that $\alpha \to \infty$, and the only reason a sharp increase may not be observed is due to the smoothing out of the derivative as T increases.

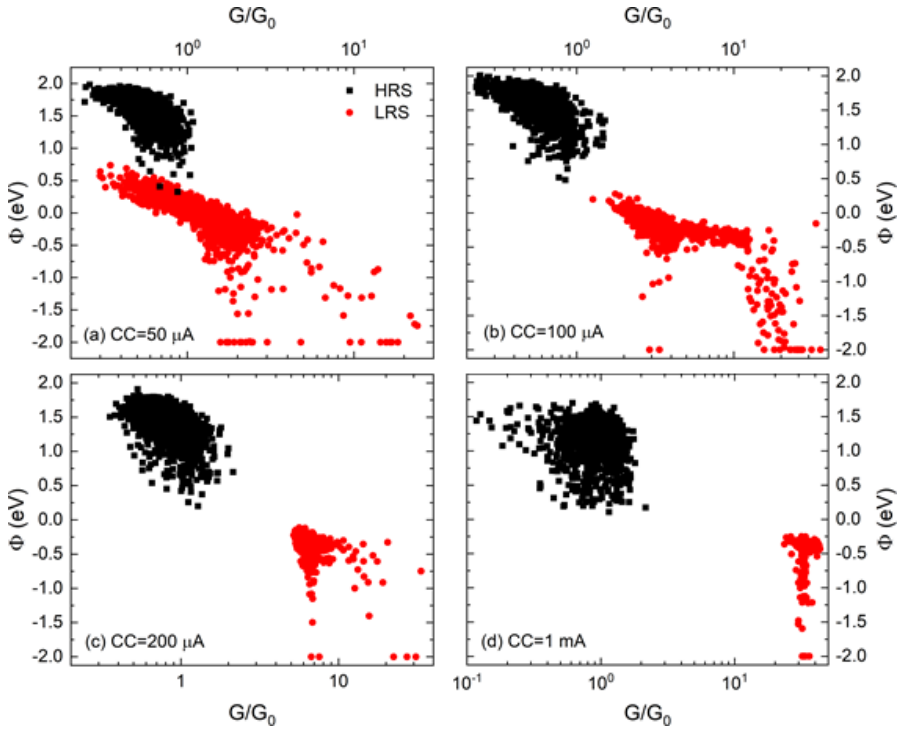

Figure 4.16 Values of Φ resulting from the application of the single subband QPC model. They are plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed.

The results of Φ (Figure 4.16) show a clear descent of the barrier height with the increase in conductance, which is consistent with the model. The descent almost appears initially exponential for the LRS in Figure 4.16 (a). The dispersion for higher values of conductance could be due to a loss of resolution resultant from the smoothing out of the derivative, for the very high transmission values reached, as discussed for α.

The correlation between α and Φ was then investigated (Figure 4.17). A clear correlation can be observed, in a consistent fashion with the one reported by Lian *et. al.* [83]: it can also be observed that αΦ is a natural parameter of the model, as demonstrated in equation (2.4) for $E = 0$. Here, the colour of the points was matched to the value of $G/G_0/N$, at an applied voltage of $0.1\ V$. The added lines were coloured according to the predicted transmission, resultant from (2.4). It can be seen that the colours match up very well. This result can be added to the range of values for both parameters, and to the simulations expressed in Figure 4.13. A conclusion can be drawn that one is indeed in the regime where modelling the overall transmission of multiple channels as $T \approx N/\{1 + \exp[\alpha(\Phi - E)]\}$ is appropriate. Thus, the physical assertion that only the first subband contributes to conduction is rejected. This is corroborated by the almost inexistence of HRS curves, and LRS curves with low conductance, with $N = 1$, which, if one was dealing with multiple individual conduction channels, would be expected.

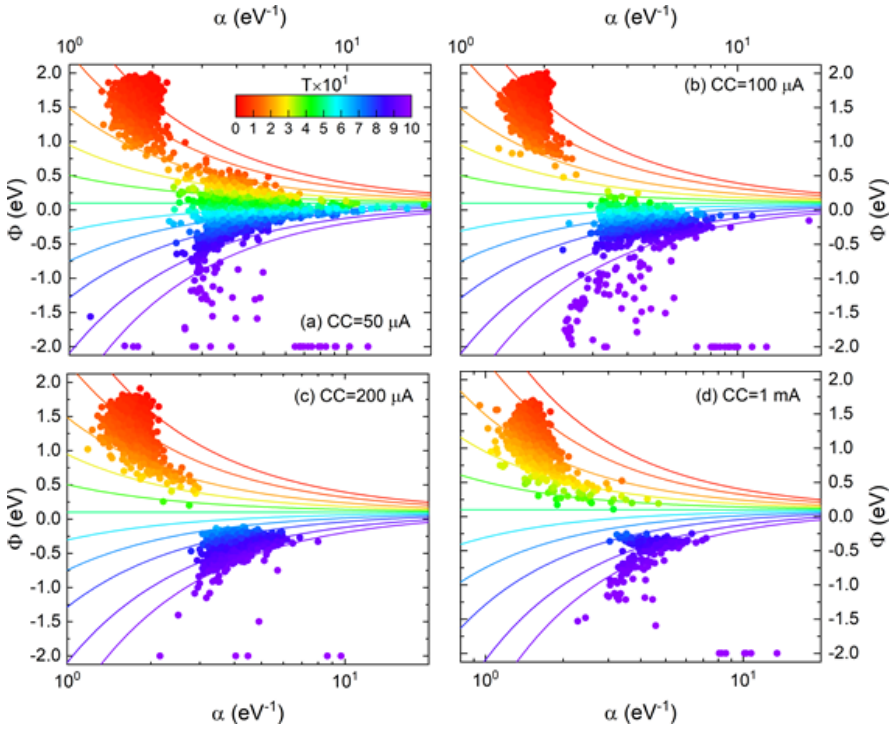

Figure 4.17 Values of Φ vs α obtained. The colour of the dots is tied to $G/G_0/N$ at 0.1 V; the colour of the lines is tied to the transmission predicted by (2.4). Both share the same colour scale.

The results for the LRS and HRS for a CC of $50\ \mu A$ also barely overlap, and are easily distinguishable despite sharing the same conductance. This means that, even with the weak RESET observed in Figure 4.4, a change in the overall device state can still be observed. One can thus conclude that this model can explain the transition between the HRS and LRS as a continuous evolution of its parameters, and even weak or failed SET operations may show a change in the overall $I - V$ curve, that with appropriate techniques may then be detected.

4.6 Results of the fixed barrier height approximation

A further approximation to the QPC model that has been suggested is to consider the case when $\Phi \gg E$ (the barrier is very high in comparison to the applied voltage), which would result in [83]

$$I \approx G_0 N \exp(-\alpha\Phi)\left(V + \frac{\alpha\beta}{2}V^2\right). \tag{4.7}$$

One can note some inconsistencies with this result, notably that the resulting transmission probability would thus read $T = \exp(-\alpha\Phi)$, which cannot be considered equivalent to the extension of (1.4), $T = [1 + \exp(\alpha\Phi)]^{-1}$. Additionally, since β is dimensionless, the term $\alpha\beta V^2/2$ has the dimensions of $[E^{-1}V^2] = [VQ^{-1}]$, needing an additional charge term to be dimensionally correct. For large values of αΦ the two expressions converge (Figure 4.6), but begin to diverge below αΦ~3. Notably, the introduced approximation is shown to not tolerate $\alpha\Phi < 0$ (meaning the barrier cannot be lower than the Fermi level), or the resultant transmission would be greater than one.

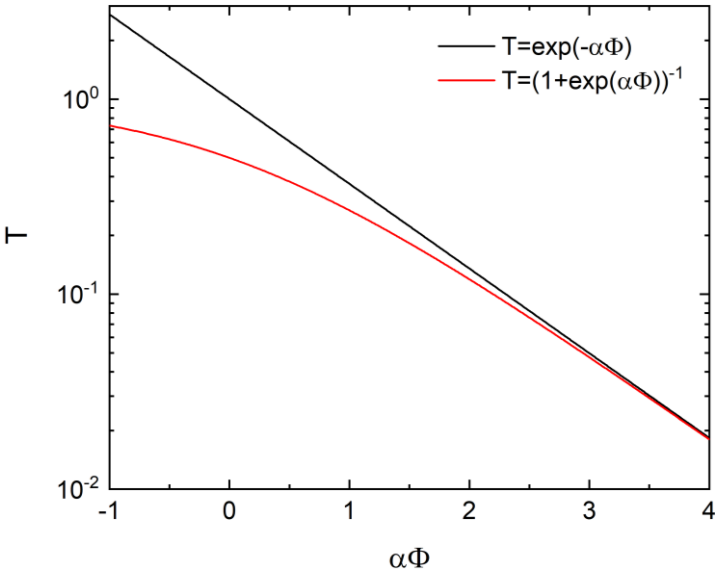

Figure 4.18 Transmission probability as a function of αΦ considering the approximation made in [83], compared to the one resultant from applying equation (2.4) with $E = 0$.

The Taylor expansion of equation (2.10) around $V = 0$ is

$$I = \frac{G_0 N}{1 + \exp(\alpha\Phi)}\left[V + \frac{e\alpha(2\beta - 1)}{2}\frac{\exp(\alpha\Phi)}{1 + \exp(\alpha\Phi)}V^2\right] + O(V^3). \tag{4.8}$$

For α=0 one also obtains a linear dependence between I and V, while for $\Phi = 0$ this is not the case. However, there are some fundamental differences between this expression and the one displayed in (2.10). The expression for the transmission obtained is indeed the one predicted by equation (2.4), and for the quadratic term the dependence on T is also quadratic. Also, given that α, as a function of ω_x, is a measure of barrier thickness (in fact, it can be shown [76] that $t_b = \hbar\alpha\sqrt{2\Phi/m\pi^2}$, where t_b is the barrier thickness measured at the Fermi level), for α=0 (an infinitely thin barrier), $T_{\alpha=0} = 0.5$, and not 1, as would be expected [80,82,136]. This results from the assumption that the potential around the bottleneck is parabolic, not accounting for the smoothing of the slope away from the constriction. One should thus consider that the disappearance of the barrier should not be considered as the limit α→0, which would require that the constriction width varied very abruptly around the bottleneck, but rather as the limit α→∞,

where the barrier is smoothed out ($\omega_x \to 0$, meaning the potential is no longer a function of x) given the disappearance of the constriction itself.

The impossibility of $\Phi < 0$, and the unity transmission expected for $\Phi = 0$, also clash with a direct interpretation of the model's origin. It requires the assumption that $V_0 \geq -\hbar\omega_y \cdot 3/2$ (from equation (2.2), for $n = 1$), which is not necessarily true, and not grounded in the model itself. This approximation is generally employed for the HRS, where the reported results contain values of Φ on the order of $1\ eV$, while the stopping voltages are of around $1\ V$ [76,78,80,82,83,136]. While the case of $\Phi < 0$ may not be relevant in that scenario, the base assumption of $\Phi \gg E$ is not valid along the entire curve, and it does not allow one to generalize the expression to the LRS.

The constant barrier height model introduced in Chapter 2 also leans on this approximation (the height is fixed, so when $\alpha = 0$ one should obtain $T = 1$), meaning the physical interpretation of the model does not hold up for thinner barriers. To appropriately model the barrier in such a fashion, more care needs to be taken in the underlying transmission expression, going for a more complex model where the parabolic approximation is cut off past a certain threshold. Alternatively, one needs to apply this approximation to very nonlinear HRS curves only (where the expectation of $\alpha\Phi$ being large leads to both transmission expressions converging). This approximation will be applied anyway, to observe how these considerations hold up to the results.

As expressed in equation (2.11), the new variables of the model are thus N and t_b, the width of the barrier measured at the Fermi level. While a normalized version of this last parameter, t_b/t_0, could be used, in this work the values proposed in [80] were used. Those are $t_0 = 1.12$ Å and $\Phi = 1.23\ eV$. Given that these were the result of their *ab initio* calculations for a TaO_x memristor, the same material being studied in this work, it is reasonable to use them as the starting point. The starting parameters used were $\left(N, t_b\ (\text{Å})\right) = (1, 2.758)$. the value for t_b was taken from the proposed formula $t_b = t_0\alpha\Phi$ for $\alpha = 2\ eV^{-1}$, to match the starting value used for the single subband model simulations presented in the previous section. The weight parameter chosen was again $\lambda^* = 1$.

The results of N (Figure 4.19) and t_b (Figure 4.20) are significantly different from those obtained in the literature [80,82] for the same model. For the HRS, a similar pattern is found, with an apparent exponential decrease of t_b and exponential increase of N with increasing G, as predicted by the model. The spread of these values is actually lower that than found in the cited papers. The results for the LRS, however, are rather distant from those predicted; in particular, the barrier width actually increases for the LRS over the HRS for all values of CC except $1\ mA$. Even on this last graph, there does not appear to be a strong correlation between t_b and G, and the barrier width actually seems to increase with increasing conductance, which is the opposite of the predicted result. The starting parameters were changed, covering lower initial values for t_b, and higher values of N, but the results found were similar. The abnormally large values of N could be an artefact of the regularization, so the weight parameter was lowered to $\lambda^* = 0.1$, finding similar results.

For the other values of CC, it is found that, for $G/G_0 > 3$, N begins to rise more slowly with growing conductance, and the barrier width starts decreasing after increasing up to this point (again, the opposite of the expected result). This odd behaviour can be explained due to the imposition of an upper limit of $N < 1000$, which was imposed to limit the trust-region search. It is when N goes above 100, and starts approaching this limit, that this phenomenon begins to manifest, and when this cap is raised it is corrected. However, obtaining such large N values is already alarming, and even the values of several hundred found in the cited papers are confusing, as these are scales that do not appear to make physical sense. The notion that this amount of parallel channels could be contributing to conduction, even if considering it a multiple subband

contribution, is rather curious. Also, even if results similar to those found in the cited papers were encountered, those were already criticized previously in this section.

Overall, the removal of a parameter from the model resulted in a poorer quality of fit and loss of physical interpretation, without gaining significant computational advantages. As such, this approximation is rejected in favour of the single subband method, which was determined to be the best form found of applying the QPC model to the acquired experimental data.

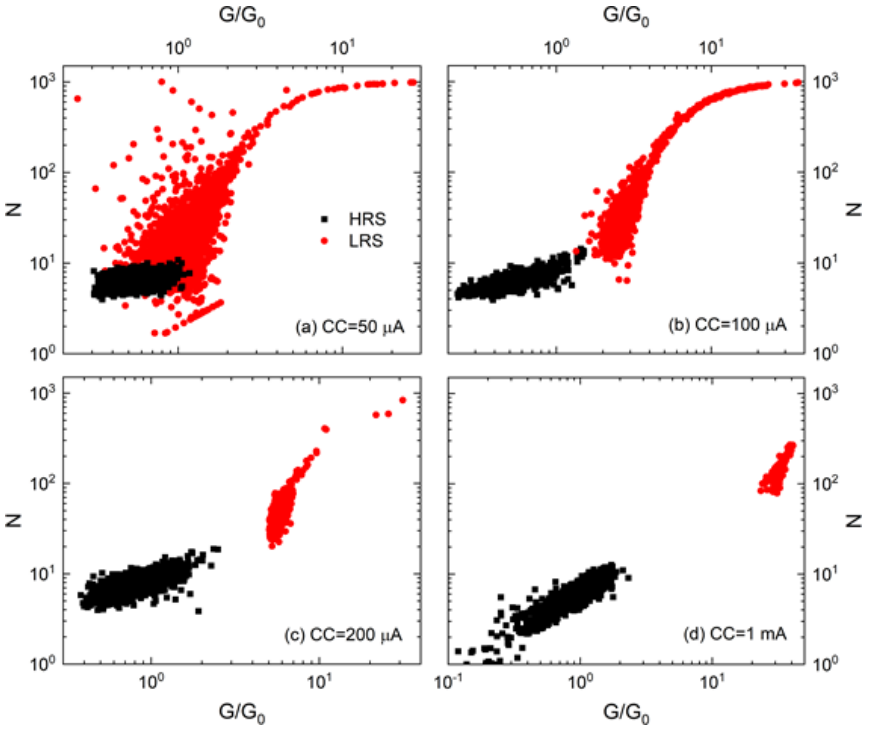

Figure 4.19 Values of N resulting from the application of the fixed barrier height approximation, plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed.

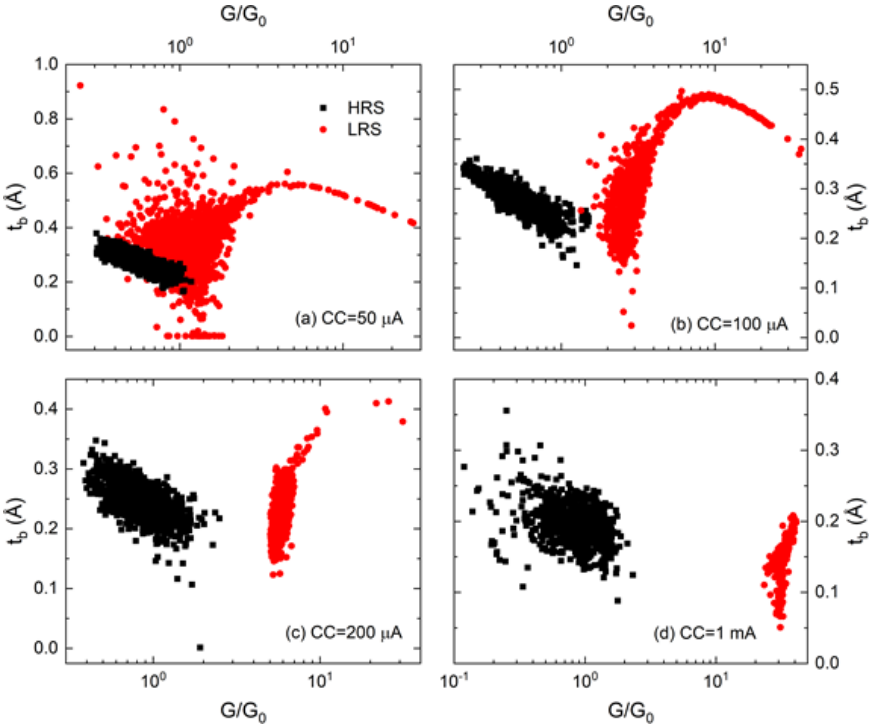

Figure 4.20 Values of t_b resulting from the application of the fixed barrier height approximation, plotted versus the normalized conductance measured at 0.1 V, for the four values of CC employed.

Conclusions and Future Work

Stable resistive switching was obtained for the Pt/Ta/Ta_2O_5/Pt samples used. The overall quality of switching improved with increasing compliance current (from $50\ \mu A$ to $1\ mA$). There, a ratio $G_{LRS}/G_{HRS} > 10$ was found, and maintained over the course of 1250 successive voltage cycles. The observed switching characteristics, and constitution of the cell, lead to the conclusion that the underlying switching mechanism is one of oxygen vacancy migration, as described by the valence change model (VCM).

Difficulties in convergence of the fitting algorithm were overcome with the use of the Levenberg-Marquardt algorithm, with a heuristic method for the determination of the weight parameter. This method was supported via leave one out cross validation (LOOCV), and showed robustness in terms of convergence and dependence on the initial guess. The single subband model was found to be more robust than the original QPC model, and easier to apply in an automated fashion due to the ability to find local minima farther away from the starting point.

The $I - V$ curves obtained for the Low (LRS) and High (HRS) Resistance States were well fitted by the Quantum Point Contact model (QPC), and it proved to be a better fit than other conventional conduction mechanisms. The observation of random telegraph signal (RTS) supports the existence of a quantization of states, which can also be explained by the QPC model upon considering a discrete variation of the filament's width.

The evolution of the model's parameters when applied to the experimental data matches expectations, showing a lowering and widening of the barrier with an increase in filament conductance. It was determined that, for higher conductances, the model's physical basis no longer holds true, but the mathematical model proved robust and capable of fitting such situations. The model can thus be applied in a general fashion, with care taken in the physical interpretations of the resulting parameters. The physical origin of the single subband model was also criticized, and the parameters should be interpreted as an effective contribution of multiple subbands in the filament.

Results of the approximation to constant barrier height do not corroborate its underlying theory. This very theory is also criticized, and the assumptions required for the approximations rejected. The notion of a quantum constriction in the CF is found to be a more sensible approach to the problem.

Overall, the first subband approximation is found to be the most reliable and efficient method of applying the QPC model to the experimental data, and the loss in physical accuracy can be overcome with a more careful interpretation of the results.

For future work, it is proposed that galvanic protection be applied to single out one individual state of the device. The influence of temperature on the switching phenomenon can then be studied with extra rigor, and the results can be interpreted from the QPC point of view. The algorithm can also be stabilized for the original interpretation of the model, and in part for the fixed barrier height approximation. If a robust and mass-applicable method of applying the original model is found, greater physical accuracy can be extracted from the model's application, and the RTS observed can be studied quantitatively. An application of the model to different materials (ZrO_2, HfO_2, …) can also be conducted, and the results compared with each other, while *ab initio* simulations can be performed to check if the assertions considered during this thesis hold true in other scenarios.

References

[1] Moore, G. E. *Electronics* **38,** 114 (1965).

[2] Moore, G. E. *Proc. Tech. Dig. Int'l Electron Devices Meet.* **21,** 11 (1975).

[3] Moore, G. E. *Proc. SPIE* **2438,** (1995).

[4] Theis, T. N. and Wong, H.-S. P. *Comput. Sci. Eng.* **19,** 41 (2017).

[5] Marinella, M. J. and Zhirnov, V. V. in *Emerging Nanoelectronic Devices* (eds. Chen, A., Hutchby, J., Zhirnov, V. & Bourianoff, G.) 246 (John Wiley & Sons Ltd., 2015).

[6] Zhirnov, V. V. and Marinella, M. J. in *Emerging Nanoelectronic Devices* (eds. Chen, A., Hutchby, J., Zhirnov, V. & Bourianoff, G.) 37 (John Wiley & Sons Ltd., 2015).

[7] Meena, J. S., Sze, S. M., Chand, U. and Tseng, T. *Nanoscale Res. Lett.* **9,** 526 (2014).

[8] Aritome, S. in *Nand Flash Memory Technologies* 129, 273 (Wiley, 2015).

[9] Wang, L., Yang, C. and Wen, J. *Electron. Mater. Lett.* **11,** 505 (2015).

[10] Menzel, S., Linn, E. and Waser, R. in *Emerging Nanoelectronic Devices* (eds. Chen, A., Hutchby, J., Zhirnov, V. & Bourianoff, G.) 137 (John Wiley & Sons Ltd., 2015).

[11] Wang, L., Yang, C., Wen, J. and Gai, S. *J. Nanomater.* **2014,** 927696 (2014).

[12] Wang, W., Pedretti, G., Milo, V., Carboni, R., Calderoni, A., Ramaswamy, N., Spinelli, A. S. and Ielmini, D. *Sci. Adv.* **4,** eaat4752 (2018).

[13] Kataeva, I., Merrikh-Bayat, F., Zamanidoost, E. and Strukov, D. in *2015 International Joint Conference on Neural Networks (IJCNN)* 1 (2015).

[14] Ielmini, D. *Microelectron. Eng.* **190,** 44 (2018).

[15] Backus, J. *Commun. ACM* **21,** 613 (1978).

[16] Chua, L. *Ieee Trans. Circuit Theory* **CT-18,** 507 (1971).

[17] Chua, L. O. and Kang, S. M. *Proc. IEEE* **64,** 209 (1976).

[18] Strukov, D. B., Snider, G. S., Stewart, D. R. and Williams, R. S. *Nature* **453,** 80–83 (2008).

[19] Meuffels, P. and Soni, R. *Cornell Univ. Libr.* arXiv:1207.7319 (2012).

[20] Di Ventra, M. and Pershin, Y. V. *Nanotechnology* **24,** 255201 (2013).

[21] Linn, E., Siemon, A., Waser, R. and Menzel, S. *IEEE Trans. Circuits Syst. I Regul. Pap.* **61,** 2402 (2014).

[22] Vongehr, S. and Meng, X. *Sci. Rep.* **5,** 11657 (2015).

[23] Sundqvist, K. M., Ferry, D. K. and Kish, L. B. *Fluct. Noise Lett.* **16,** 1771001 (2017).

[24] Abraham, I. *Sci. Rep.* **8,** 10972 (2018).

[25] Waser, R. in *Memristive Phenomena - From Fundamental Physics to Neuromorphic Computing* (eds. Waser, R. & Wuttig, M.) 15 (Forschungszentrum Jülich GmbH, 2016).

[26] Schroeder, H., Zhirnov, V. V., Cavin, R. K. and Waser, R. *J. Appl. Phys.* **107,** 054517 (2010).

[27] Hickmott, T. W. *J. Appl. Phys.* **33,** 2669 (1962).

[28] Huang, R., Tang, Y., Kuang, Y., Ding, W., Zhang, L. and Wang, Y. *IEEE Trans. Electron Devices* **59,** 3578–3582 (2012).

[29] Huang, R., Cai, Y., Liu, Y., Bai, W., Kuang, Y. and Wang, Y. in *Proceedings - IEEE International Symposium on Circuits and Systems* 838 (2014).

[30] Wagner, P. J. and Estabrook, G. F. *Adv. Electron. Mater.* **1,** 1400003 (2015).

[31] Liu, Y., Li, F., Chen, Z., Guo, T., Wu, C. and Kim, T. W. *Vacuum* **130,** 109–112 (2016).

[32] Shan, Y., Lyu, Z., Guan, X., Younis, A., Yuan, G., Wang, J., Li, S. and Wu, T. *Phys. Chem. Chem. Phys.* **20,** 23837 (2018).

[33] Yamamoto, S., Kitanaka, T., Miyashita, T. and Mitsuishi, M. *Nanotechnology* **29,** 26LT02 (2018).

[34] Guiot, V., Cario, L., Janod, E., Corraze, B., Phuoc, V. T., Rozenberg, M., Stoliar, P., Cren, T. and Roditchev, D. *Nat. Commun.* **4,** 1172 (2013).

[35] Chu, D., Lin, X., Younis, A., Li, C. M., Dang, F. and Li, S. *J. Solid State Chem.* **214,** 38 (2014).

[36] Zheng, M., Ni, H., Xu, X., Qi, Y., Li, X. and Gao, J. *Phys. Rev. Appl.* **9,** 044039 (2018).

[37] Jeong, D. S., Schroeder, H., Breuer, U. and Waser, R. *J. Appl. Phys.* **104,**

123716 (2008).
[38] Gomez-Marlasca, F., Ghenzi, N., Rozenberg, M. J. and Levy, P. *Appl. Phys. Lett.* **98,** 042901 (2011).
[39] Borgatti, F., Park, C., Herpers, A., Offi, F., Egoavil, R., Yamashita, Y., Yang, A., Kobata, M., Kobayashi, K., Verbeeck, J., Panaccione, G. and Dittmann, R. *Nanoscale* **5,** 3954 (2013).
[40] Fowler, B. W., Chang, Y. F., Zhou, F., Wang, Y., Chen, P. Y., Xue, F., Chen, Y. T., Bringhurst, B., Pozder, S. and Lee, J. C. *RSC Adv.* **5,** 21215 (2015).
[41] Sharma, A. A., Karpov, I. V., Kotlyar, R., Kwon, J., Skowronski, M. and Bain, J. A. *J. Appl. Phys.* **118,** 114903 (2015).
[42] Nagashima, K., Yanagida, T., Oka, K., Taniguchi, M., Kawai, T., Kim, J. S. and Park, B. H. *Nano Lett.* **10,** 1359 (2010).
[43] Yao, Y. P., Liu, Y. K., Dong, S. N., Yin, Y. W., Yang, S. W. and Li, X. G. *Appl. Phys. Lett.* **100,** 193504 (2012).
[44] Prakash, A., Park, J. S., Song, J., Lim, S. J., Park, J. H., Woo, J., Cha, E. and Hwang, H. in *4th International Symposium on Next-Generation Electronics, IEEE ISNE 2015* **27,** 59–60 (2015).
[45] Sandouk, E. J., Gimzewski, J. K. and Stieg, A. Z. *Sci. Technol. Adv. Mater.* **16,** 045004 (2015).
[46] Jung, J., Kwon, D., Jung, H., Lee, K., Yoon, T. S., Kang, C. J. and Lee, H. H. *J. Ind. Eng. Chem.* **64,** 85 (2018).
[47] Schindler, C., Thermadam, S. C. P., Waser, R. and Kozicki, M. N. *IEEE Trans. Electron Devices* **54,** 2762 (2007).
[48] Yanagida, T., Nagashima, K., Oka, K., Kanai, M., Klamchuen, A., Park, B. H. and Kawai, T. *Sci. Rep.* **3,** 1657 (2013).
[49] Choi, D. and Soo Kim, C. *Appl. Phys. Lett.* **104,** 193507 (2014).
[50] Ismail, M., Ahmed, E., Rana, A. M., Talib, I. and Nadeem, M. Y. *J. Alloys Compd.* **646,** 662 (2015).
[51] Santos, D. A. A., Zeng, H. and Macêdo, M. A. *Mater. Res. Bull.* **66,** 147 (2015).
[52] Zhao, X., Fan, Z., Xu, H., Wang, Z., Xu, J., Ma, J. and Liu, Y. *J. Mater. Chem. C* **6,** 7195 (2018).
[53] Yang, Y., Sheridan, P. and Lu, W. *Appl. Phys. Lett.* **100,** 203112 (2012).
[54] Chai, Y., Wu, Y., Takei, K., Chen, H.-Y., Yu, S., Chan, P. C. H., Javey, A. and Wong, H.-S. P. *IEEE Trans. Electron Devices* **58,** 3933 (2011).
[55] Gao, S., Zeng, F., Wang, M., Wang, G., Song, C. and Pan, F. *Sci. Rep.* **5,** 15467 (2015).
[56] Li, S., Wei, X., Lei, Y., Yuan, X. and Zeng, H. *Appl. Surf. Sci.* **389,** 977 (2016).
[57] Chen, X., Hu, W., Li, Y., Wu, S. and Bao, D. *Appl. Phys. Lett.* **108,** 053504 (2016).
[58] Smith, J., Chung, S., Jang, J., Biaou, C. and Subramanian, V. *IEEE Trans. Electron Devices* **64,** 4310 (2017).
[59] Sassine, G., La Barbera, S., Najjari, N., Minvielle, M., Dubourdieu, C. and Alibart, F. *J. Vac. Sci. Technol. B, Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.* **34,** 012202 (2016).
[60] Valov, I., Waser, R., Jameson, J. R. and Kozicki, M. N. *Nanotechnology* **22,** 254003 (2011).
[61] Kozicki, M. N., Mitkova, M. and Valov, I. in *Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications* (eds. Ielmini, D. & Waser, R.) 483 (Wiley, 2016).
[62] Menzel, S., Klopstra, B., Kügeler, C., Böttger, U., Staikov, G. and Waser, R. *MRS Proc.* **1160,** 1160-H09-03 (2009).
[63] Zhuge, F., Li, K., Fu, B., Zhang, H., Li, J., Chen, H., Liang, L., Gao, J., Cao, H., Liu, Z. and Luo, H. *AIP Adv.* **5,** 057125 (2015).
[64] Ielmini, D., Bruchhaus, R. and Waser, R. *Phase Transitions* **84,** 570 (2011).
[65] Kim, J., Jung, K., Kim, Y., Jo, Y., Cho, S., Woo, H., Lee, S., Inamdar, A. I., Hong, J., Lee, J. K., Kim, H. and Im, H. *Sci. Rep.* **6,** 23930 (2016).
[66] Waser, R. *J. Nanosci. Nanotechnol.* **12,** 7628 (2012).
[67] Magyari-Köpe, B., Tendulkar, M., Park, S. G., Lee, H. D. and Nishi, Y. *Nanotechnology* **22,** 254029 (2011).
[68] Joshua Yang, J., Miao, F., Pickett, M. D., Ohlberg, D. A. A., Stewart, D. R., Lau, C. N. and Williams, R. S. *Nanotechnology* **20,** 215201 (2009).
[69] Yang, J. J., Strachan, J. P., Miao, F., Zhang, M. X., Pickett, M. D., Yi, W., Ohlberg, D. A. A., Medeiros-Ribeiro, G. and Williams, R. S. *Appl. Phys. A Mater. Sci. Process.* **102,** 785 (2011).
[70] Lim, E. and Ismail, R. *Electronics* **4,** 586 (2015).
[71] Chiu, F.-C. *Adv. Mater. Sci. Eng.* **2014,**

1 (2014).
[72] Kao, K. C. in *Dielectris Phenomena in Solids* 327, 381 (Elsevir, 2004).
[73] Simmons, J. G. *J. Appl. Phys.* **34,** 1793 (1963).
[74] Scott, J. F. *J. Phys. Condens. Matter* **26,** 142202 (2014).
[75] Jiang, J., Lim, D. G., Ramadoss, K. and Ramanathan, S. *Solid. State. Electron.* **146,** 13–20 (2018).
[76] Miranda, E. and Suñé, J. in *2001 IEEE International Reliability Physics Symposium Proceedings* 367 (2001).
[77] Belmonte, A., Celano, U., Redolfi, A., Fantini, A., Muller, R., Vandervorst, W., Houssa, M., Jurczak, M. and Goux, L. *IEEE Trans. Electron Devices* **62,** 2007 (2015).
[78] Miranda, E. A., Walczyk, C., Wenger, C. and Schroeder, T. *IEEE Electron Device Lett.* **31,** 609 (2010).
[79] Degraeve, R., Fantini, A., Raghavan, N., Goux, L., Clima, S., Chen, Y. Y., Belmonte, A., Cosemans, S., Govoreanu, B., Wouters, D. J., Roussel, P., Kar, G. S., Groeseneken, G. and Jurczak, M. in *Proceedings of the 21th International Symposium on the Physical and Failure Analysis of Integrated Circuits* 245 (2014).
[80] Lian, X., Wang, M., Rao, M., Yan, P., Yang, J. J. and Miao, F. *Appl. Phys. Lett.* **110,** 173504 (2017).
[81] van Ruitenbeek, J., Masis, M. M. and Miranda, E. in *Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications* (eds. Ielmini, D. & Waser, R.) 197 (Wiley, 2016).
[82] Lian, X., Cartoixà, X., Miranda, E., Perniola, L., Rurali, R., Long, S., Liu, M. and Suñé, J. *J. Appl. Phys.* **115,** 244507 (2014).
[83] Lian, X., Long, S., Cagli, C., Buckley, J., Miranda, E., Liu, M. and Suñe, J. in *2012 13th International Conference on Ultimate Integration on Silicon* 101 (2012).
[84] Stathis, J. H. *J. Appl. Phys.* **86,** 5757 (1999).
[85] Degraeve, R., Groeseneken, G., Bellens, R., Depas, M. and Maes, H. E. *Proc. Int. Electron Devices Meet.* 863 (1995).
[86] Wharam, D. A., Thornton, T. J., Newbury, R., Pepper, M., Ahmed, H., Frost, J. E. F., Hasko, D. G., Peacock, D. C., Ritchie, D. A. and Jones, G. A. C. *J. Phys. C Solid State Phys.* **21,** L209 (1988).
[87] Van Wees, B. J., Van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwenhoven, L. P., Van Der Marel, D. and Foxon, C. T. *Phys. Rev. Lett.* **60,** 848 (1988).
[88] Landauer, R. *IBM J. Res. Dev.* **1,** 223 (1957).
[89] Büttiker, M. *Phys. Rev. Lett.* **57,** 1761 (1986).
[90] Büttiker, M. *Phys. Rev. B - Condens. Matter Mater. Phys.* **41,** 7906 (1990).
[91] Tekman, E. and Ciraci, S. *Phys. Rev. B - Condens. Matter Mater. Phys.* **43,** 7145 (1991).
[92] Degraeve, R., Roussel, P., Goux, L., Wouters, D., Kittl, J., Altimime, L., Jurczak, M. and Groeseneken, G. in *2010 International Electron Devices Meeting* 28.4.1 (2010).
[93] Connor, J. N. L. *Mol. Phys.* **15,** 37–46 (1968).
[94] Miller, W. H. *J. Chem. Phys.* **48,** 1651 (1968).
[95] Rosário, C. M. M., Thöner, B., Schönhals, A., Menzel, S., Wuttig, M., Waser, R., Sobolev, N. A. and Wouters, D. J. *Appl. Phys. Lett.* **112,** 213504 (2018).
[96] Zhong, X., Rungger, I., Zapol, P., Nakamura, H., Asai, Y. and Heinonen, O. *Phys. Chem. Chem. Phys.* **18,** 7502–7510 (2016).
[97] Goux, L., Fantini, A., Chen, Y. Y., Redolfi, A., Degraeve, R. and Jurczak, M. *ECS Solid State Lett.* **3,** Q79 (2014).
[98] Tsurumaki-Fukuchi, A., Nakagawa, R., Arita, M. and Takahashi, Y. *ACS Appl. Mater. Interfaces* **10,** 5609 (2018).
[99] Guo, Y. and Robertson, J. *Appl. Phys. Lett.* **105,** 223516 (2014).
[100] Keithley. in *Series 2600A System SourceMeter Reference Manual* (2013).
[101] Keithley. in *6th Edition Low Level Measurements Handbook* 3.16 (2004).
[102] Avriel, M. in *Nonlinear Programming: Analysis and Methods* 288 (Dover Publications, 2003).
[103] Gill, P. E., Murray, W. and Wright, M. H. in *Practical Optimization* 83, 133 (Academic Press, 1981).
[104] Kelley, C. T. in *Iterative Methods for Optimization* 13, 39, 71 (SIAM Frontiers in Applied Mathematics, no 18, 1999).
[105] Gill, P. E. and Murray, W. *SIAM J.*

Numer. Anal. **15,** 977 (1978).
[106] Pujol, J. *Geophysics* **72,** W1 (2007).
[107] Levenberg, K. *Q. Appl. Math.* **2,** 164 (1943).
[108] Marquardt, D. W. *J. Soc. Ind. Appl. Math.* **11,** 431 (1963).
[109] Kuhn, M. and Johnson, K. in *Applied Predictive Modeling* 61 (Springer, 2013).
[110] Golub, G. H., Heath, M. and Wahba, G. *Technometrics* **21,** 215 (1979).
[111] Chen, A. in *2012 IEEE International Reliability Physics Symposium Proceedings* MY.2.1 (2012).
[112] Wan, H. J., Zhou, P., Ye, L., Lin, Y. Y., Tang, T. A., Wu, H. M. and Chi, M. H. *IEEE Electron Device Lett.* **31,** 246 (2010).
[113] Ambrogio, S., Milo, V., Wang, Z. Q., Balatti, S. and Ielmini, D. *IEEE Electron Device Lett.* **37,** 1268 (2016).
[114] Shrestha, P., Nminibapiel, D., Campbell, J. P., Cheung, K. P., Baumgart, H., Deora, S. and Bersuker, G. in *2013 IEEE International Integrated Reliability Workshop Final Report* 55 (2013).
[115] Luo, Y., Shao, S., Hu, H., Li, J., Shen, J. and Zhao, D. *Mater. Lett.* **201,** 169 (2017).
[116] Mickel, P. R., Lohn, A. J. and Marinella, M. J. *Appl. Phys. Lett.* **105,** 053503 (2014).
[117] Schönhals, A., Rosário, C. M. M., Hoffmann-Eifert, S., Waser, R., Menzel, S. and Wouters, D. J. *Adv. Electron. Mater.* **4,** 1700243 (2018).
[118] Gries, U. N., Schraknepper, H., Skaja, K., Gunkel, F., Hoffmann-Eifert, S., Waser, R. and De Souza, R. A. *Phys. Chem. Chem. Phys.* **20,** 989 (2018).
[119] Menzel, S., Waters, M., Marchewka, A., B??ttger, U., Dittmann, R. and Waser, R. *Adv. Funct. Mater.* **21,** 4487 (2011).
[120] Spiess, A. N. and Neumeyer, N. *BMC Pharmacol.* **10,** 1–11 (2010).
[121] Colin Cameron, A. and Windmeijer, F. A. G. *J. Econom.* **77,** 329–342 (1997).
[122] Yeargan, J. R. and Taylor, H. L. *J. Appl. Phys.* **39,** 5600–5604 (1968).
[123] Aygun, G. and Turan, R. *Thin Solid Films* **517,** 994–999 (2008).
[124] Young, L. *Proc. R. Soc. London A* **244,** 41 (1958).
[125] Wilcox, P. S. and Westwood, W. D. *Can. J. Phys* **49,** 1543 (1971).
[126] Husted, D., Gruss, L. and Mackus, T. *J. Electrochem. Soc.* **118,** 1989 (1971).
[127] Smith, D. J. and Young, L. *IEEE Trans. Electron Devices* **28,** 22 (1981).
[128] Kukli, K., Ritala, M. and Leskelä, M. *J. Electrochem. Soc.* **142,** 1670 (1995).
[129] Devine, R. A. B., Vallier, L., Autran, J. L., Paillet, P. and Leray, J. L. *Appl. Phys. Lett.* **68,** 1775 (1996).
[130] Inoue, N., Ozaki, T., Monnaka, T., Kashiwabara, S. and Fujimoto, R. *Jpn. J. Appl. Phys* **36,** 704 (1997).
[131] Zaima, S., Furuta, T., Yasuda, Y. and Ida, M. *J. Electrochem. Soc.* **137,** 1297 (1990).
[132] Lau, W. S., Qian, P. W., Sandler, N. P., McKinley, K. A. and Chu, P. K. *Jpn. J. Appl. Phys* **36,** 661 (1997).
[133] Nakagawa, Y. and Okada, T. *J. Appl. Phys.* **68,** 556 (1990).
[134] Zaima, S., Furuta, T., Koide, Y. and Yasudo, Y. *J. Electrochem. Sac.* **137,** 2876 (1990).
[135] Kukli, K., Aarik, J., Aidla, A., Kohan, O., Uustare, T. and Sammelselg, V. *Thin Solid Films* **260,** 135 (1995).
[136] Walczyk, C., Walczyk, D., Schroeder, T., Bertaud, T., Sowińska, M., Lukosius, M., Fraschke, M., Wolansky, D., Tillack, B., Miranda, E. and Wenger, C. *IEEE Trans. Electron Devices* **58,** 3124 (2011).
[137] Hung, K. K., Ko, P. K., Hu, C. and Cheng, Y. C. *IEEE Electron Device Letters* **11,** 90 (1990).
[138] Simoen, E., Fang, W., Aoulaiche, M., Luo, J., Zhao, C. and Claeys, C. *Thin Solid Films* **613,** 2 (2016).
[139] Ranjan, A., Raghavan, N., Shubhakar, K., Thamankar, R., Molina, J., O'Shea, S. J., Bosman, M. and Pey, K. L. in *2016 IEEE International Reliability Physics Symposium (IRPS)* 7A–4–1 (2016).
[140] Moré, J. J. and Sorensen, D. C. *SIAM J. Sci. Stat. Comput.* **4,** 553 (1983).
[141] Cao, X., Li, X. M., Gao, X. D., Zhang, Y. W., Liu, X. J., Wang, Q. and Chen, L. D. *Appl. Phys. A Mater. Sci. Process.* **97,** 883 (2009).
[142] Sun, B., Liu, L., Xu, N., Gao, B., Wang, Y., Han, D., Liu, X., Han, R. and Kang, J. *Jpn. J. Appl. Phys.* **48,** 04C061 (2009).
[143] Wu, Y. L., Liao, C. W. and Ling, J. J. *Appl. Phys. Lett.* **104,** 242906 (2014).